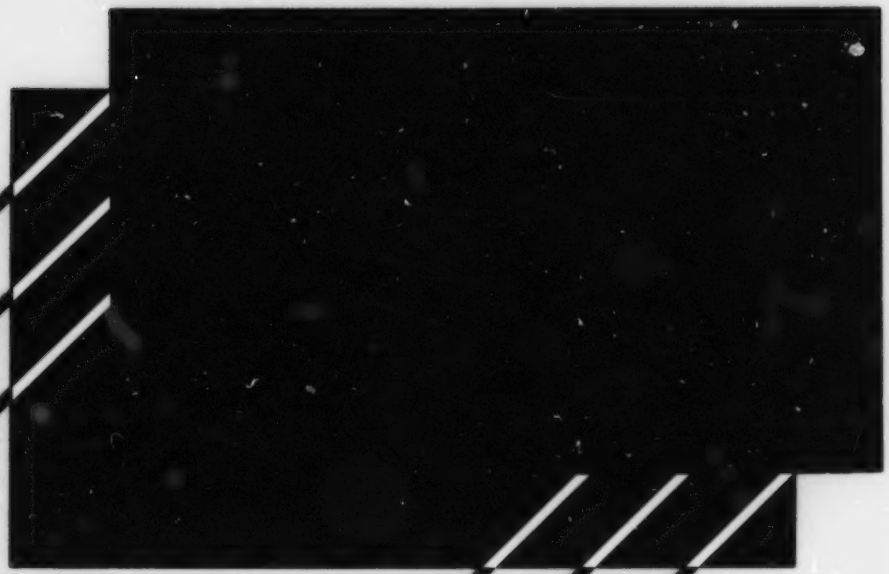
Gouvernement du Canada
Ministère des Communications

Government of Canada
Department of Communications

Le Centre canadien de recherche sur l'informatisation du travail
Canadian Workplace Automation Research Centre



Canada

IRMA

UN INTERPRÈTE DE RAPPORTS DE MARCHÉS AGRICOLES

Jean-Luc Cochard
Pierre Isabelle
Michel Simard

Groupe de traduction assistée

Centre canadien de recherches sur
l'informatisation du travail
Ministère des Communications

Laval

mai 1990

92-05715/

II

Nº de cat. Co 28-1/49-1990 F
ISBN 0-662-96152-8

Les opinions émises dans ce rapport n'engagent que leurs auteurs.

* This report is also available in English.

Remerciements

Nous tenons à exprimer notre gratitude envers

Le CCRIT qui, par son financement, a rendu possible cette réalisation.

Raymond Descout, responsable du groupe de technologies vocales, qui a lancé l'idée du projet IRMA et qui s'est personnellement impliqué dans l'organisation et les relations avec Expotec.

Russ Short, pour sa participation à la programmation de l'interface graphique d'IRMA.

Elliott Macklovitch, pour sa participation à l'élaboration du scénario et pour ses précieux commentaires lors de la rédaction de ce rapport.

Pierre Hamel, François Brochu, Richard Cheng et Philippe Galipeau pour la réalisation des logiciels de communication entre machines.

Michel Beaudet et Pierre Brazeau pour leur diligence à résoudre les problèmes matériels et logiciels.

Marc Dymetman, d'avoir accepté de lire et de commenter ce rapport.

Tous les autres membres du groupe de traduction assistée pour les remarques et suggestions lors des différentes phases du développement d'IRMA.

L'ensemble des démonstrateurs du kiosque IRMA et tous les responsables d'Expotec qui ont cru à ce projet.

Et finalement, Sun Canada, Quintus-Prolog et Zenith Canada qui ont gracieusement fourni le matériel et le logiciel essentiels à la réalisation de ce projet.

Table des matières

Résumé.....1

1. Introduction.....3
1.1. L'historique du projet IRMA.....3
1.2. L'objectif didactique et le contexte d'utilisation.....4
1.3. La traduction assistée par ordinateur et le système CRITTER.....5
1.4. Le traitement de la parole.....5
1.5. L'interface d'utilisation.....6

2. Les possibilités d'IRMA.....7
2.1. Les généralités.....7
Le kiosque IRMA.....7
Le scénario général d'utilisation.....8
Le niveau de départ d'IRMA.....8
2.2. Le tableau de commandes.....9
Les paramètres d'exécution.....9
Les paramètres de configuration.....11
La commande de traduction.....12
Les portes de sortie.....13
2.3. L'éditeur vocal.....14
Les types d'unités affichées.....15
Les manipulations d'unités.....15
Le clavier et la poubelle.....16
L'effet secondaire de la traduction.....17
2.4. L'affichage des traductions.....18
2.5. Les étapes de la traduction.....18

3. Les idées directrices.....27
3.1. Le profil des démonstrateurs et le public-cible.....27
3.2. Les thèmes majeurs.....28
Les difficultés de la TAO.....28
La perte d'informations de la langue parlée par rapport à la langue écrite.....28
La puissance de CRITTER en matière de désambiguïsation.....29
3.3. L'organisation de l'interface.....29
Les principes de clarté et de sobriété.....29

Le langage des symboles graphiques 30
Le flux des informations et la symétrie des affichages 31
Les actions-réactions explicites 32
L'utilisation simple de la souris 33
3.4. La conception des outils 33
La centralisation des commandes des périphériques 33
Le jeu restreint de commandes de l'éditeur vocal 33

4. Les détails techniques 35
4.1. Le groupe de Contrôle et de Traitement 36
4.2. Le groupe de Reconnaissance 37
Le système VoiceScribe-1000 et le logiciel DragonKEY 37
L'environnement VoiceScribe pour IRMA 38
Les communications entre les groupes de Reconnaissance et de Contrôle 39
4.3. Le groupe de Synthèse 41
DECtalk 42
TELEVOX 42
Les communications à l'intérieur du groupe de Synthèse et avec le groupe de Contrôle 42

5. Les variantes d'IRMA 45
5.1. L'expérience Tubach 45
5.2. IRMA Multi-processus 46

Conclusion 49

Bibliographie 51

Annexe 53

Résumé

IRMA est le nom d'un système informatique qui reproduit les compétences d'un interprète humain. Il permet ainsi de reconnaître un texte dicté à un microphone, de traduire ce texte de l'anglais vers le français ou inversement, dépendant de la langue source et, finalement, de fournir une version orale de la traduction.

Tout ce processus est réalisée grâce à la mise bout à bout de systèmes commerciaux de *reconnaissance* et de *synthèse vocales* et d'un système expérimental pour la partie *traduction automatique* – le système CRITTER.

La comparaison entre IRMA et un interprète humain ne va pas au-delà d'une simple similitude de fonction exercée par les deux. L'intérêt d'IRMA réside surtout dans sa capacité d'illustrer, grâce à une interface très élaborée, le fonctionnement de ses mécanismes internes.

Ce rapport décrit en détail chacune des composantes du système et les principes sous-jacents à la conception de l'interface d'utilisation.

Domaines d'intérêts: Traduction automatique, reconnaissance vocale, synthèse vocale, interfaces graphiques, Prolog.

Introduction

1

Le système informatique que nous présentons ici a comme nom IRMA, ce qui est l'abréviation de INTERPRÈTE DE RAPPORTS DE MARCHÉS AGRICOLES. Le sens que nous conférons ici à "interprète" est bien celui de traducteur simultané entre deux langues. De même qu'un interprète humain, IRMA est un système informatique qui reconnait des phrases prononcées en français ou en anglais, qui les traduit dans l'autre langue officielle et fournit une forme orale de cette traduction.

Cependant la comparaison entre IRMA et l'interprète humain ne doit pas être amenée au-delà de cette équivalence fonctionnelle. IRMA n'est qu'un système expérimental qui ne peut pas rivaliser avec un humain dans le domaine de l'interprétation. Chacun des maillons qui le compose est beaucoup plus faible que les compétences humaines équivalentes. IRMA illustre simplement les possibilités actuelles dans des domaines variés de l'intelligence artificielle que sont la reconnaissance et la synthèse de la parole et la traduction assistée par ordinateur (TAO).

La figure 1 décrit schématiquement le système IRMA. À gauche de la figure, le microphone permet la saisie vocale et symbolise le système de reconnaissance vocale. À droite, le haut-parleur qui symbolise la synthèse vocale, émet des sons. Entre les deux, recevant des données du microphone et transmettant des résultats au haut-parleur, figure le système CRITTER de traduction automatique, schématisé par une "boîte noire".

Figure 1 Schéma du système informatique IRMA.

1.1. L'historique du projet IRMA

IRMA est né d'un projet conjoint du groupe de Traduction Assistée par Ordinateur (TAO) qui fait partie de la Direction des Technologies Avancées (DLT) et du groupe des Technologies Vocales (GTV) faisant partie de la Direction des Systèmes Intégrés (DLD), deux groupes du Centre canadien de recherches sur l'informatisation du travail (CCRIT).

Le moteur de cette réalisation est une proposition de participation à Expotec 89[1], une exposition estivale et thématique de vulgarisation scientifique qui se tient au Vieux-Port de Montréal. Le thème de l'été 1989 était: "Zoom sur les Sens!". L'exposition dressait donc un panorama des technologies et des phénomènes scientifiques reliés aux sens de la vue, de l'ouïe et de la parole.

La collaboration entre le GTV et le groupe de TAO a consisté, pour le premier, en un apport de compétences dans les domaines de la reconnaissance et de la synthèse de la parole et en des développements de routines de communication et de gestion des périphériques vocaux et, pour le second, en une mise à disposition du système CRITTER et en un développement d'une interface d'utilisation du logiciel d'IRMA.

Entre la prise de décision d'une participation à Expotec 89 et la première démonstration officielle et complète du poste, un peu plus de deux mois seulement se sont écoulés durant lesquels il a fallu tester les différents périphériques vocaux, inventer des scénarios de démonstration, concevoir et implanter l'interface d'utilisation, raccorder ensemble les différentes composantes matérielles, définir et mettre en place les différents protocoles de communication et adapter le système CRITTER à certaines particularités du projet IRMA.

1.2. L'objectif didactique et le contexte d'utilisation

Le principal défi que nous avons voulu relever en réalisant cette application, est celui de la vulgarisation scientifique. IRMA se veut un outil didactique qui illustre les difficultés inhérentes à l'automatisation de la traduction. Cette illustration est d'autant plus difficile que le spectateur moyen n'envisage pas à quel point cette tâche peut être complexe et approximative. Pour lui, la traduction humaine est perçue comme une opération atomique, indivisible dont il a peine à imaginer la somme de connaissances qu'elle nécessite.

Le problème est tout à fait similaire avec la reconnaissance et la synthèse vocales, deux domaines dans lesquels l'être humain excelle et dans lesquels, comparativement, l'ordinateur balbutie. La mise en situation de ce genre de périphériques dans IRMA offre donc la possibilité au spectateur de juger des performances des technologies actuelles dans ces domaines. Le système de reconnaissance vocale utilisé dans IRMA, en particulier, ne correspond pas à la fine pointe de la technologie mais il en fournit une bonne approximation. Malgré cela, le fonctionnement des deux périphériques vocaux, de reconnaissance et de synthèse, est très spectaculaire et a l'avantage de capter l'attention ou mieux de l'entretenir durant les explications plus techniques et ardues sur la TAO.

Si la vulgarisation scientifique doit affiner le sens critique des personnes auxquelles elle s'adresse, il n'est pas évident que ce but soit atteint dans un cadre comme Expotec. Le nombre de kiosques à visiter et le temps moyen accordé à chacun, plus proche d'une minute que de cinq, ne favorise pas la réflexion en profondeur sur un sujet aussi technique que la TAO. Néanmoins, si quelqu'un prend la peine de s'attarder à IRMA, nous sommes convaincu que le potentiel didactique de la démonstration permet de répondre à la plus grande partie de ses interrogations.

---

1 Présentation de Lavalin Communications et le Vieux-Port de Montréal

1.3. La traduction assistée par ordinateur et le système CRITTER

La construction d'une machine qui traduirait des textes fut une des premières applications non numériques suggérées pour les ordinateurs lors de leur apparition vers la fin des années 40. Pourquoi donc, après quarante années d'investissements et d'efforts énormes, la traduction automatique et de grande qualité de textes quelconques est-elle toujours un but inaccessible? La raison en est que la simple substitution mot à mot n'est pas suffisante pour produire de bonnes traductions. Les ordinateurs, comme les traducteurs humains, ont besoin de comprendre très précisément le texte à traduire. Notre attitude face à ce problème revient à affirmer que la traduction par ordinateur ne fournit des résultats de qualité que si elle s'applique à des domaines très restreints. C'est pourquoi IRMA ne traduit que des rapports de marchés agricoles: une connaissance profonde du domaine a été introduite dans le système.

Comme le laisse voir la figure 1, la phase de traduction proprement dite est prise en charge par le système CRITTER qui est un logiciel de traduction assistée par ordinateur en cours de réalisation au CCRIT [1]. Ce logiciel n'a pas été développé dans le but de servir de système de traduction à IRMA. Mais il faut plutôt voir IRMA comme une application possible s'appuyant sur les compétences de CRITTER. Dans son état actuel, le système CRITTER offre la possibilité de traduire automatiquement de l'anglais au français et inversement des textes extraits d'un corpus réel: des rapports de marché agricole émis par Agriculture Canada.

En tant que composante centrale d'IRMA, le système CRITTER a été légèrement modifié afin de traiter un problème de taille pour lequel il n'était pas conçu à priori: la désambiguïsation des homophones hétérographes (cf. section 3.2). Ce problème qui apparaît lors de l'utilisation du périphérique de reconnaissance vocale est généralement difficile à gérer et la prise en charge satisfaisante qu'y en est faite par CRITTER est un heureux imprévu.

1.4. Le traitement de la parole

La reconnaissance de la parole est une autre tâche par laquelle l'ordinateur essaie de se rapprocher de l'être humain. La compétence humaine que l'on essaie ainsi de reproduire est faussement simple puisqu'il ne s'agit que d'identifier des mots prononcés. La complexité apparaît lorsqu'on se rend compte qu'on fait automatiquement des liaisons lorsqu'on prononce une séquence de mots, que les intonations de différents locuteurs sont forcément différentes, que l'environnement d'écoute est rarement idéal, etc.

Actuellement, les systèmes commerciaux les plus répandus offrent la possibilité de reconnaître des mots isolés qui ont été clairement prononcés par des locuteurs pour lesquels le système a subi un entraînement. Ceci illustre bien les limitations de ce genre d'application et ça explique aussi la raison pour laquelle le vocabulaire d'IRMA est si limité: chaque utilisateur ou utilisatrice du système doit effectuer un entraînement durant lequel il ou elle doit répéter très soigneusement quatre ou cinq fois chaque mot du vocabulaire d'IRMA.

La synthèse de la parole pose à priori moins de problèmes que la reconnaissance et elle est en conséquence beaucoup moins contraignante. Il n'en demeure pas moins que pour atteindre une qualité "humaine" d'élocution, la synthèse vocale devient une tâche hautement

complexe. IRMA intègre deux systèmes de synthèse vocale, un pour l'anglais et un pour le français dont les qualités respectives d'élocution ne sont malheureusement pas comparables.

1.5. L'interface d'utilisation

Toutes ces composantes du système IRMA: le système CRITTER et les périphériques vocaux, sont connectées entre elles au sein d'une *interface d'utilisation*. Cette interface, en plus de gérer le flux de données entre ces diverses composantes, présente visuellement, de manière sobre et claire, les différentes fonctions disponibles dans l'application. C'est donc elle que nous allons principalement décrire dans les chapitres suivants non pas parce qu'il s'agit d'un élément de liaison mais plutôt parce que l'interface donne une image du produit, parce qu'elle concrétise les intentions des réalisateurs et qu'elle joue un rôle primordial dans le processus de vulgarisation.

Ainsi donc, l'organisation de la console d'IRMA et la description des fonctions disponibles font l'objet du chapitre suivant: *Les possibilités d'IRMA* (chapitre 2). L'apparence visuelle de l'interface d'utilisation d'IRMA répond à une série de critères et de principes que nous avons préalablement discutés et qui sont condensés dans le chapitre 3: *Les idées directrices de la conception*. Afin de compléter notre panorama d'IRMA, nous abordons ensuite quelques détails très techniques dans le chapitre 4: *Les détails techniques*, comme le matériel nécessaire au fonctionnement d'IRMA, les protocoles de communication entre l'interface et les périphériques, etc.

IRMA a aussi été un banc d'essai pour l'utilisation d'autres systèmes de reconnaissance vocale que celui retenu pour Expotec. Ce développement est résumé dans le chapitre 5: *Les variantes d'IRMA*. En conclusion, dans le chapitre 6, nous proposons une réflexion critique sur ce que l'on pourrait appeler l'expérience ou l'aventure IRMA, où nous essayons de faire ressortir les erreurs, les mauvais choix mais aussi les échos positifs et les répercussions d'IRMA sur le travail futur du groupe en matière d'interfaces graphiques. Comme répercussion concrète, on notera la mise sur pied d'un projet conjoint avec la compagnie Bourbeau Pinard, Inc. pour la création d'un Dictionnaire Micro-informatisé du Français Phonétique (DMFP), en tant qu'extension du DMF [2].

Pour compléter ce survol relativement exhaustif du projet IRMA, nous proposons, en annexes, d'une part, un document distribué aux démonstrateurs pour les guider dans la démonstration du poste IRMA: le scénario de démonstration et, d'autre part, une série de copies d'écran de la console d'IRMA qui illustre ce scénario et donne une idée du fonctionnement du système.

Les possibilités d'IRMA

2

2.1. Les généralités

Le kiosque IRMA

La figure 2 représente le kiosque IRMA, tel que présenté durant tout l'été à Expotec 89. Les différentes composantes visibles d'IRMA sont les suivantes:

- une station SUN 4/110, avec son clavier et la souris, à laquelle nous faisons référence en parlant de *la console d'IRMA*,
- un micro-casque sur la tête du démonstrateur pour la reconnaissance vocale,
- un haut-parleur pour la synthèse vocale,
- un amplificateur encastré dans le meuble.

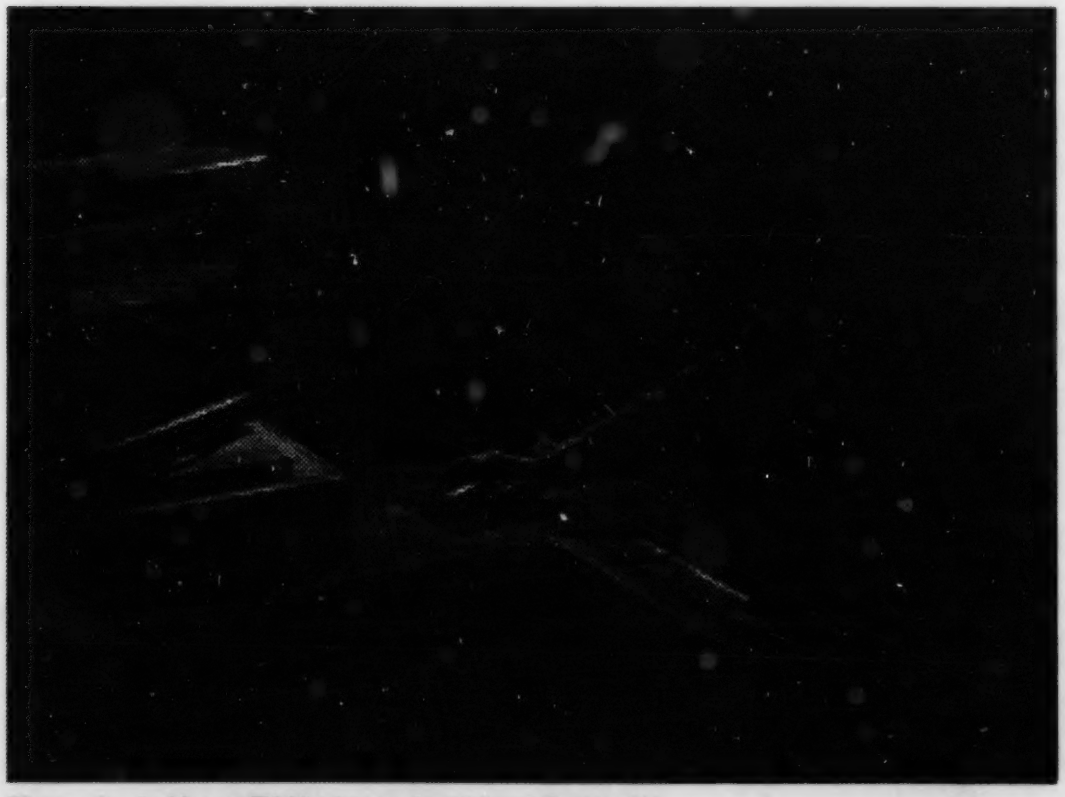


Figure 2 Le kiosque IRMA tel que présenté durant Expotec 89.

Le meuble contient toute une gamme de matériels divers dont nous préciserons la nature et le rôle dans le chapitre 4. Accroché à l'arrière du meuble, apparaît un tableau explicatif qui résume très sommairement les compétences d'IRMA.

Dans la suite de ce chapitre, nous allons concentrer notre attention exclusivement sur le contenu de la console d'IRMA, puisque l'interaction avec le démonstrateur passe essentiellement par ce périphérique.

Le scénario général d'utilisation

Avant d'entrer dans le détail des possibilités de l'interface, voici une description sommaire du scénario typique d'utilisation d'IRMA. Le présentateur soumet des phrases à IRMA en prenant garde que chaque mot de la phrase fasse partie du vocabulaire du système. Les phrases sont prononcées, mot par mot, dans un micro-casque qui est branché au système de reconnaissance vocale. Lorsqu'un mot est reconnu, il est automatiquement affiché à l'écran de la console d'IRMA et c'est alors que le mot suivant peut être prononcé.

Une fois que la phrase est complètement reconnue, le présentateur active la commande de traduction. Il peut alors choisir de montrer ou non des résultats intermédiaires du processus de traduction. En fin de traduction, le système affiche automatiquement une version écrite du texte traduit et fournit parallèlement une version orale.

Pour autant qu'on s'abstienne d'afficher les résultats intermédiaires, le processus complet de traduction d'une phrase incluant la reconnaissance et la synthèse vocale ne prend pas plus de quelques secondes.

Le niveau de départ d'IRMA

Lors de la mise en marche d'IRMA, l'écran de la console a l'apparence de la figure 3. Cet écran est décomposé en plusieurs régions appelées *les fenêtres* et délimitées soit par un double trait fin, soit par un simple trait épais. Chaque fenêtre a un rôle spécifique et contient divers objets de type textuel ou graphique, appelé *un icone* dans le deuxième cas. Toutes les composantes sont détaillées dans les sections suivantes. Nous allons nous contenter pour l'instant d'en faire le survol.

La fenêtre dans la partie supérieure de l'écran qui contient le texte `Texte source :`[1] et les deux icones: le clavier et la poubelle, est nommée l'*éditeur vocal*. C'est dans cette fenêtre qu'apparaît le texte à traduire, qui est éventuellement dicté au microphone. La fenêtre dans la partie inférieure de l'écran qui contient le texte `Texte cible :` et l'icone haut-parleur, sert à l'*affichage des traductions*. Entre ces deux fenêtres, on trouve deux autres fenêtres. Celle qui ne contient rien par défaut joue un rôle dans l'affichage des *étapes de la traduction*. L'autre fenêtre qui contient le texte INTERFACE DE DÉMO IRMA et quantité d'autres objets est le *tableau de commandes* centralisé de l'interface. C'est depuis cette fenêtre que l'on commande une traduction, que l'on branche ou débranche les périphériques de reconnaissance et de synthèse vocales, etc.

Remarques

Dans le contexte du bilinguisme officiel au Canada, il est important qu'IRMA soit en mesure d'afficher des objets textuels en français ou en anglais. C'est un des paramètres du

---

1 Ce jeu de caractères est utilisé systématiquement lorsqu'on cite le contenu d'un objet présent à la console d'IRMA.

système (cf. *Les paramètres de configuration*, dans la section 2.2) qui permet de sélectionner la langue d'affichage. On peut d'ailleurs faire remarquer que la prise en charge de cette particularité a nécessité une certaine rigueur de programmation qui permettrait d'augmenter aisément la liste des langues d'affichage.

D'ailleurs, la version anglaise de ce rapport s'appuie sur les copies d'écran de la version anglaise d'IRMA.

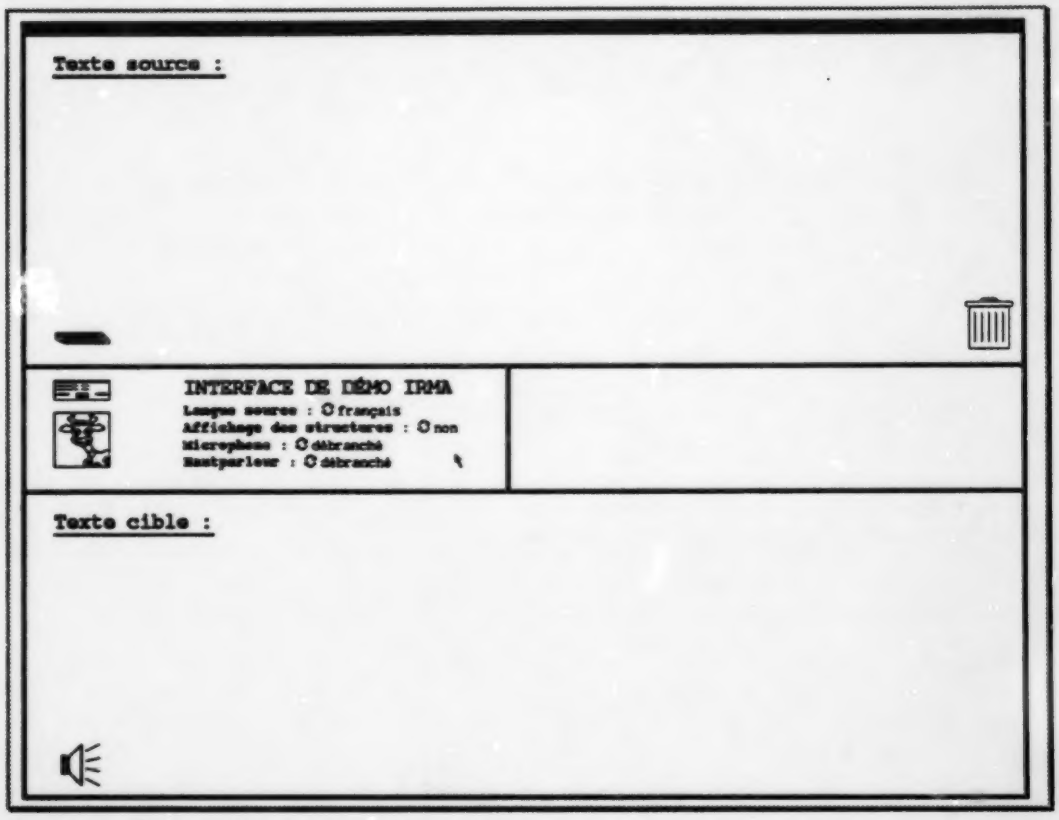


Figure 3 Écran d'IRMA, dans la situation de départ.

2.2. Le tableau de commandes

Le tableau de commandes d'IRMA permet d'exécuter trois types d'opérations: changer l'état des paramètres de fonctionnement, commander une traduction ou quitter le programme. Les paramètres de fonctionnement peuvent être catalogués en deux catégories: ceux que l'on peut vouloir modifier à tout instant durant une démonstration et que nous nommons donc *paramètres d'exécution*, et ceux qui sont modifiés accessoirement entre les démonstrations, appelés les *paramètres de configuration*.

Les paramètres d'exécution

Nous avons retenu quatre paramètres d'exécution qui apparaissent directement dans le tableau de commandes nommés: **`Langue source`**, **`Affichage des structures`**, **`Microphone`** et **`Haut-parleur`** (cf. figure 4).

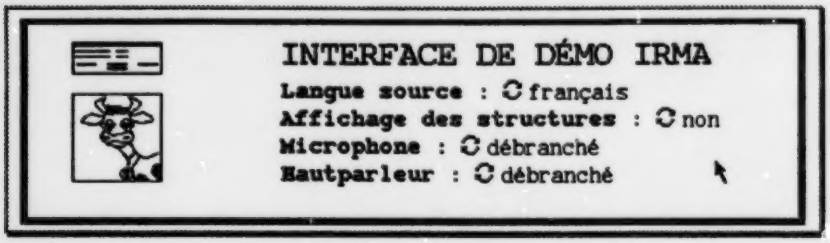
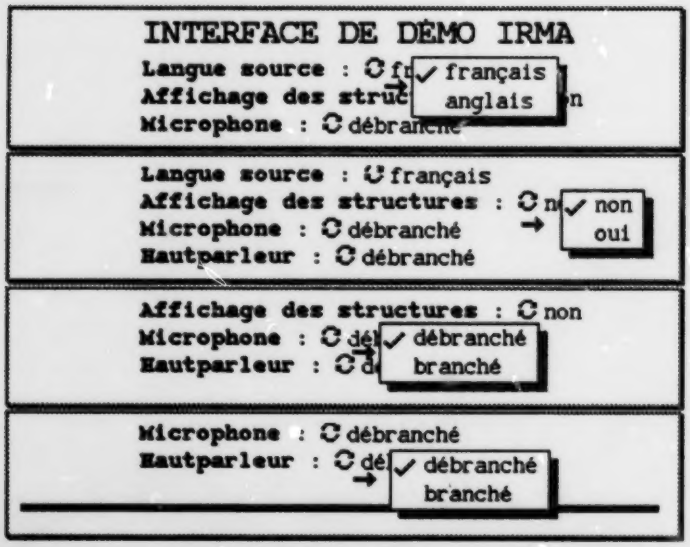


Figure 4 Fenêtre contenant le tableau de commandes d'IRMA.

Chacun des paramètres peut être modifié de deux manières différentes: soit en appuyant sur le bouton de gauche de la souris, en ayant le curseur à proximité de la valeur courante, soit en appuyant sur le bouton de droite. Dans le premier cas, la valeur courante est remplacée par la valeur suivante dans la liste des valeurs disponibles et ceci de façon cyclique (une fois la fin de la liste atteinte, on recommence avec la première valeur possible). Dans le deuxième cas, le système affiche la liste des possibilités, en mettant en évidence la valeur courante. Nous avons opté pour cette technique afin d'exhiber les valeurs possibles des différents paramètres.



Figures 5.1 à 5.4 Les différentes valeurs possibles des paramètres d'exécution.

Le paramètre **Langue source** (cf. figure 5.1) fait référence à la langue dans laquelle est fourni le texte apparaissant dans l'éditeur vocal. Ce paramètre influence la reconnaissance vocale, puisque la détermination de la langue parlée est nécessaire pour sélectionner le bon vocabulaire (cf. section 4.2), mais elle influence aussi de manière évidente le processus de traduction.

Le paramètre **Affichage des structures** (cf. figure 5.2) est en relation directe avec la commande de traduction (cf. *La commande de traduction*, dans la section 2.2). La valeur **non** signifie que la traduction est effectuée sans interruption, ni visualisation des structures manipulées durant ce processus. La valeur **oui** enclenche un mode de

fonctionnement de la commande de traduction qui permet d'expliquer le fonctionnement du système CRITTER (cf. section 2.5) qui réalise la traduction au sein du système IRMA.

Le paramètre **Microphone** (cf. figure 5.3) détermine le statut du périphérique de reconnaissance vocale, à savoir s'il est **branché** ou **débranché**. Ceci permet donc au démonstrateur d'alterner entre des explications et la dictée d'une phrase à traduire.

Le paramètre **Haut-parleur** et non "Hautparleur"! (cf. figure 5.4) détermine le statut du périphérique de synthèse vocale. Si le haut-parleur est **branché**, le texte traduit est transmis à la synthèse vocale après être affiché dans la fenêtre d'affichage des traductions. S'il est **débranché**, seul l'affichage à l'écran a lieu.

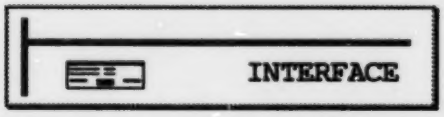


Figure 6 Icone permettant l'affichage des paramètres de configuration.

Les paramètres de configuration

La présentation des paramètres de configuration se fait à partir du tableau de commandes, en sélectionnant l'icone qui schématise une version réduite d'une fenêtre (cf. figure 6), à l'aide du bouton de gauche de la souris. Lorsque cette sélection est effectuée, le système fait apparaître une nouvelle fenêtre temporaire au-dessus du tableau de commandes et le curseur est automatiquement placé à l'intérieur de celle-ci (cf. figure 7). Pour faire disparaître cette fenêtre temporaire, il suffit de déplacer le curseur hors de ses limites.

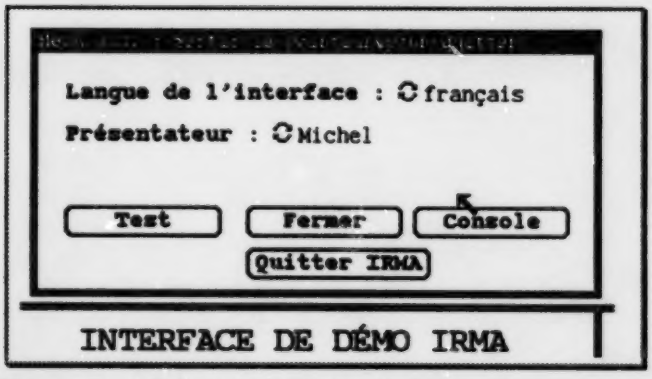
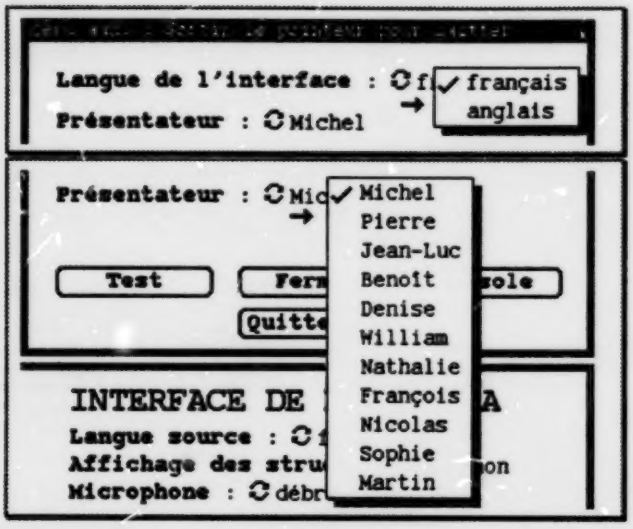


Figure 7 Affichage des paramètres de configuration.

Le paramètre **Langue de l'interface** (cf. figure 8.1) fait référence à la langue dans laquelle sont affichés les objets textuels de l'interface, c'est-à-dire les différents titres, les noms des paramètres et leurs valeurs et les messages qui apparaissent accessoirement à l'écran dans le déroulement de certaines opérations. C'est à l'aide de ce paramètre qu'IRMA peut être configuré pour des anglophones ou des francophones.



Figures 8.1 et 8.2 Les différentes valeurs possibles des paramètres de configuration.

Le paramètre **Présentateur** (cf. figure 8.2) permet de configurer optimalement la reconnaissance vocale. Dans la mesure où le logiciel de reconnaissance vocale intégré à IRMA est mono-locuteur, ce paramètre s'impose d'office.

La commande de traduction

Le processus de traduction est activé en sélectionnant l'icone de la vache avec le bouton gauche de la souris. Ainsi le contenu de l'éditeur vocal est soumis à une traduction par le système CRITTER.

Lorsque l'utilisateur soumet une phrase au système CRITTER, l'icone de la vache change d'aspect. La vache passe par une série d'attitudes distinctes qui permettent de suivre l'évolution du processus de traduction. Voici une description qui met en parallèle l'attitude et le traitement en cours.

État de départ, en attente d'une demande de traduction (icone par défaut).

En phase de traitement linguistique, exécution d'une étape du système CRITTER.

En phase de traitement dans l'interface, soit affichage de la traduction, soit calcul ou affichage des structures graphiques.

Échec de l'opération en cours; cet icone s'affiche pendant environ deux secondes avant de laisser place à l'icone par défaut.

Grâce à ce codage imagé, une traduction par la système CRITTER correspond, pour la vache, à manger et à avaler une gerbe de foin. Si la traduction échoue, l'air dépité de l'animal peut être interprété comme une indigestion!

La commande de traduction a deux effets distincts dans le cas d'une traduction sans affichage de structures. Le premier effet évident est de fournir une phrase traduite. Le second effet est de désambiguïser la phrase originelle. L'utilisation d'un périphérique de reconnaissance vocale pour la saisie du texte source engendre des ambiguïtés lexicales dues aux homophones hétérographes. La représentation de ces ambiguïtés est décrite dans la section 2.3. Un effet secondaire de la traduction est de devoir choisir parmi toutes les possibilités affichées, laquelle est la bonne. En conséquence, la phrase originelle correcte peut être mise en évidence à l'intérieur de l'éditeur vocal.

Les portes de sortie

Le système IRMA dispose de trois "portes de sortie", qui correspondent à trois situations particulières dans lesquelles l'utilisateur désire quitter IRMA (cf. figure 9).

Figure 9 Fenêtre contenant les commandes de sortie d'IRMA.

La première situation, la plus simple, est celle où l'utilisateur désire quitter définitivement le programme IRMA. C'est le bouton **`Quitter IRMA`** qui est associé à cette action. Dans ce cas, IRMA demande une confirmation en affichant une nouvelle fenêtre (cf. figure 10) et en permettant de choisir entre les réponses **`Oui`**, **`Non`** à la question **`Quitter ?`**.

Figure 10 Demande de confirmation lors de l'exécution de la commande **`Quitter IRMA`**.

IRMA offre aussi la possibilité d'interrompre le travail dans IRMA en libérant l'écran pour l'affichage d'autres fenêtres. Cette opération s'exécute en activant le bouton **`Fermer`**. L'écran d'IRMA est ainsi remplacé par un icone contenant la vache (cf. figure 11). Pour réinstaller l'écran d'IRMA, il suffit de sélectionner l'icone avec le bouton gauche de la souris.

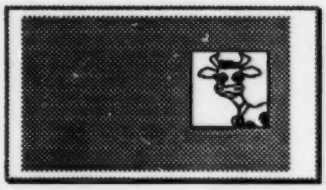


Figure 11 Portion d'écran après la fermeture temporaire de l'application IRMA.

La troisième possibilité offerte par le bouton `Console` permet de stopper temporairement IRMA en donnant le contrôle au PC de reconnaissance afin d'entraîner la reconnaissance de nouveaux mots. Cette opération très particulière est commentée plus largement dans la section 4.2. Dans ce cas de sortie, IRMA demande aussi une confirmation par l'intermédiaire d'une nouvelle fenêtre (cf. figure 12), qui offre les réponses `Oui`, `Non` à la question `Le contrôle va au PC. OK?`

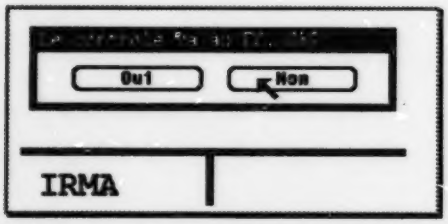


Figure 12 Demande de confirmation lors de l'exécution de la commande **Console**.

Le rôle du dernier bouton `Test` est documenté dans la section 4.2.

2.3. L'éditeur vocal

L'éditeur vocal est le nom donné à la fenêtre qui affiche le texte reconnu par le périphérique de reconnaissance vocale (cf. figure 13). Le texte est composé d'unités séparées que sont les mots et les signes de ponctuation. En fait, il s'agit des unités déterminées par la reconnaissance vocale. En plus de l'affichage, il est possible d'effectuer des manipulations d'unités, ce qui justifie le nom d'éditeur donné à cette fenêtre.

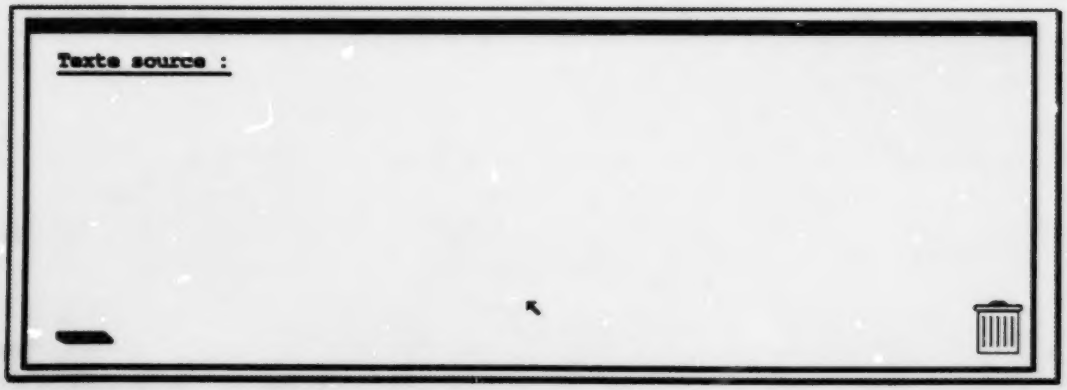


Figure 13 L'éditeur vocal d'IRMA.

Les types d'unités affichées

Il existe deux types d'unités affichables par l'éditeur vocal: les mots (incluant les signes de ponctuation) et les homophones hétérographes (cf. *La perte d'informations de la langue parlée par rapport à la langue écrite*, dans la section 3.2). Les mots sont affichés conventionnellement, les homophones sont affichés en colonne sur un support distinct (cf. figure 14).

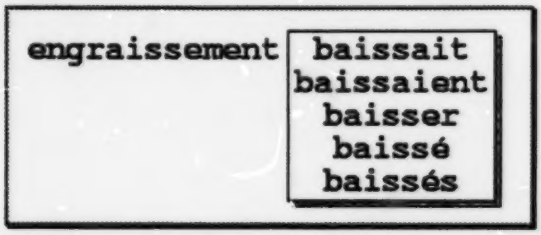


Figure 14 Portion de phrase mettant en valeur les deux types d'unités affichables.

Les manipulations d'unités

Il existe trois manipulations d'unités possibles dans l'éditeur vocal: l'adjonction d'une nouvelle unité, le remplacement d'une unité et l'élimination d'une unité. Pour effectuer ces manipulations il existe une seule opération élémentaire appelée la sélection d'une unité.

La sélection d'une unité s'effectue avec le bouton gauche de la souris en ayant le curseur à proximité ou sur l'unité. La sélection est alors symbolisée par une inversion du contraste de l'unité (cf. figure 15). Pour désélectionner une unité, il suffit d'effectuer la sélection d'une autre unité ou bien il suffit de sélectionner "dans le vide".

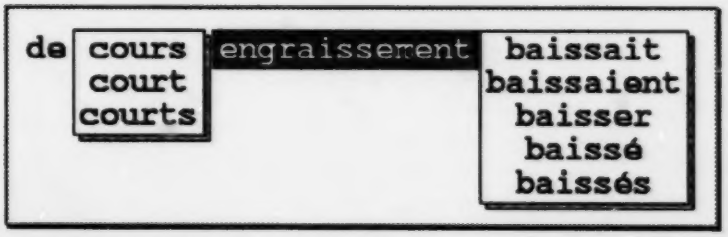


Figure 15 Affichage d'une unité sélectionnée, le mot `engraissement`.

Si aucune sélection n'est effectuée, ce qui est la situation de défaut, la dictée d'un mot est interprétée comme une adjonction d'une unité en fin de ligne. Si une sélection est visible, la dictée d'un mot est interprétée comme un remplacement de l'unité sélectionnée. Si une unité est sélectionnée deux fois d'affilée, l'unité en question est éliminée (cf. figure 16). Ces quelques opérations élémentaires constituent l'ensemble des manipulations d'unités disponibles dans l'éditeur vocal.

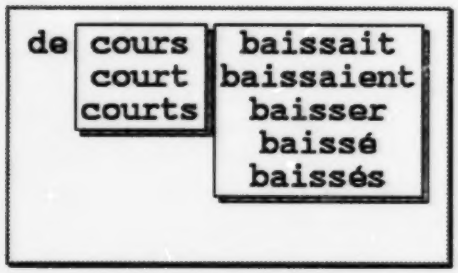


Figure 16 Illustration de l'élimination par la double sélection du mot **engraissement**.

Le clavier et la poubelle

L'éditeur vocal offre une alternative à la dictée vocale pour la saisie d'unités. Il s'agit du clavier. Par défaut, le clavier n'est pas actif dans l'éditeur vocal. Il ne suffit donc pas de taper des mots au clavier pour les faire apparaître à l'écran. Il faut commencer par activer le clavier en sélectionnant l'icone correspondant (cf. figure 17).

Figure 17 Icone clavier appartenant à l'éditeur vocal.

Lorsque cette sélection est effectuée, l'icone est temporairement caché par une fenêtre qui affichera le texte en cours de frappe au clavier, unité par unité (cf. figure 18). Lors de l'ouverture de cette fenêtre, le curseur est automatiquement placé à l'intérieur de celle-ci ce qui permet de débuter immédiatement la saisie au clavier. Lorsqu'une fin d'unité est rencontrée dans la séquence des caractères tapés au clavier – un espace entre deux mots ou un signe de ponctuation – l'unité est transmise à l'éditeur vocal comme si elle provenait du périphérique de reconnaissance vocale. Pour terminer la saisie au clavier, il suffit de quitter la fenêtre de saisie avec le curseur.

Figure 18 Fenêtre temporaire de saisie de texte au clavier.

Remarque

Lorsque la saisie de texte se fait par le clavier, il n'existe aucune possibilité de faire apparaître des homophones, puisqu'il s'agit d'une particularité de la langue parlée.

Afin d'accélérer le rythme des démonstrations d'IRMA et dans la mesure où l'utilisation de la reconnaissance vocale n'est constante durant toute la démonstration, nous avons préenregistré quelques phrases. Lorsque la **langue source** est le **français**, les phrases associées aux touches de fonction **F1** à **F4** sont les suivantes:

F1 les prix augmentaient de 3 dollars lundi, demeurant soutenus pour le reste de la semaine.
F2 les vaches grasses se vendaient à rabais, alors que les meilleures génisses augmentaient de 3 à 4 dollars.
F3 les sujets maigres de un an se vendaient à prix soutenus.
F4 les cours des bovins de court engraissement baissaient de cinq dollars.

Lorsque la **langue source** est l'**anglais**, les phrases préenregistrées sont les suivantes:

F1 prices increased three dollars Monday, remaining steady for the rest of the week.
F2 fat cows were discounted, while the top heifers were up $3-$4.
F3 lean yearlings sold at steady prices.
F4 feader prices decreased $5.

Lorsque le traitement d'une phrase est terminé et que l'on désire passer à une autre phrase, ou bien lorsque le contenu de l'éditeur vocal est trop erroné pour que des manipulations d'unités soit envisageable, on peut sélectionner l'icone poubelle (cf. figure 19). L'opération associée à cet icone consiste à nettoyer le contenu de l'éditeur vocal et de la fenêtre d'affichage des traductions (cf. section 2.4).

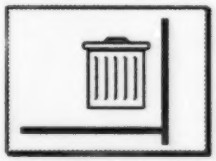


Figure 19 Icone poubelle appartenant à l'éditeur vocal.

L'effet secondaire de la traduction

Lorsqu'une traduction est réalisée au sein d'IRMA, des choix s'opèrent obligatoirement parmi les différents groupes d'homophones éventuellement présents dans la phrase source. Afin d'illustrer ces choix, le système remplace les groupes d'homophones sur leurs supports distincts par des unités conventionnelles. Ce remplacement s'effectue par un déplacement visible à l'écran des supports des homophones. Ce déplacement laisse ainsi apparaître une des unités qui apparaissait dans la liste précédente (cf. figures 15 et 20).

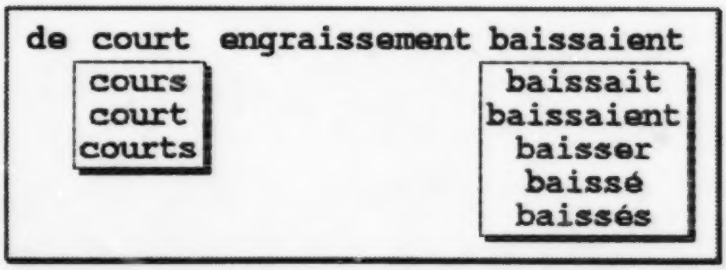


Figure 20 Représentation d'une désambiguisation d'une portion de phrase.

Cette opération est effectuée soit avant de fournir la traduction lorsque l'**affichage des structures** est inhibé, soit à la fin de l'analyse lorsque l'**affichage des structures** est activé.

2.4. L'affichage des traductions

La fenêtre d'affichage des traductions ne sert qu'à présenter la version écrite de la traduction réalisée par CRITTER (cf. figure 21). Suivant le statut du haut-parleur, la phrase est ou n'est pas envoyée au périphérique de synthèse vocale.

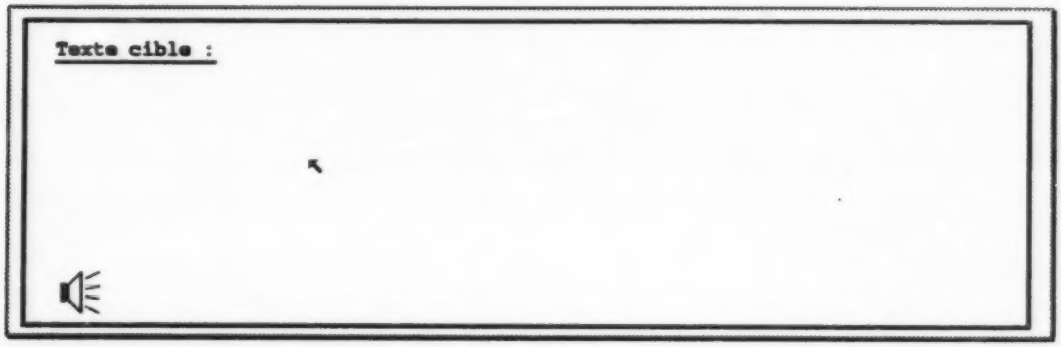


Figure 21 La fenêtre d'affichage des traductions.

Cette fenêtre contient un icone haut-parleur qui permet de réeffectuer la synthèse vocale de la phrase contenue dans la fenêtre, pour autant que le **haut-parleur** soit préalablement **branché**. Aucune autre opération n'est possible à l'intérieur de cette fenêtre puisque le nettoyage du contenu de la fenêtre est asservi à l'icone poubelle de l'éditeur vocal et qu'aucune manipulation du texte n'est prévue.

2.5. Les étapes de la traduction

Lorsqu'on active l'**affichage des structures** dans le tableau de commandes, la fenêtre à droite du tableau de commandes laisse apparaître trois nouveaux icones contenant les initiales des termes-clés pour ce mode de fonctionnement: analyse, transfert et synthèse (cf. figure 22).

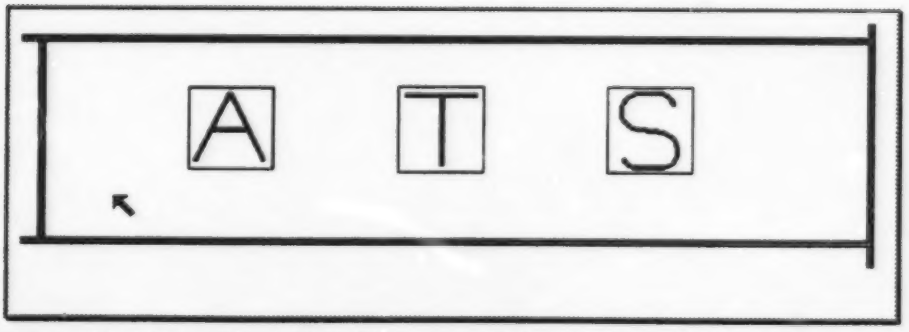


Figure 22 Configuration initiale des icones symbolisant les étapes de la traduction.

Cette possibilité d'IRMA sert à illustrer les principes et les techniques mis en oeuvre dans la réalisation d'un système de TAO comme le système CRITTER. La décomposition schématique du processus de traduction en trois étapes (cf. figure 23) met en évidence quelques grands principes:

- L'analyse et la synthèse sont des opérations dépendant d'un seul paramètre, la langue pour laquelle elles sont construites; seul le transfert est défini pour une paire de langues.
- Chaque étape ou opération est un programme qui fournit un résultat qui devient la donnée de l'étape suivante, ce qui autorise une réalisation totalement indépendante des trois étapes.
- L'analyse correspond à la phase de lecture et de compréhension du texte à traduire. Cette opération utilise les ressources d'un dictionnaire monolingue et une grammaire de la langue source. La structure sémantique, résultat de cette étape, est une représentation formelle du sens de la phrase analysée qui décrit les objets apparaissant dans la phrase de même que leurs relations implicites ou explicites à l'intérieur de celle-ci.
- Le transfert sert à produire à partir du graphe sémantique de la phrase analysée un graphe sémantique équivalent, mais orienté vers la langue cible. C'est l'étape proprement dit de traduction, qui utilise essentiellement un dictionnaire bilingue de correspondances entre structures sémantiques.
- La synthèse effectue la composition d'une phrase dans la langue cible à partir du graphe sémantique cible. À cette fin, le système va consulter le dictionnaire monolingue et les règles de la grammaire de la langue cible, de façon à ordonner correctement les mots, conjuguer les verbes, effectuer les accords de genre et nombre, etc.

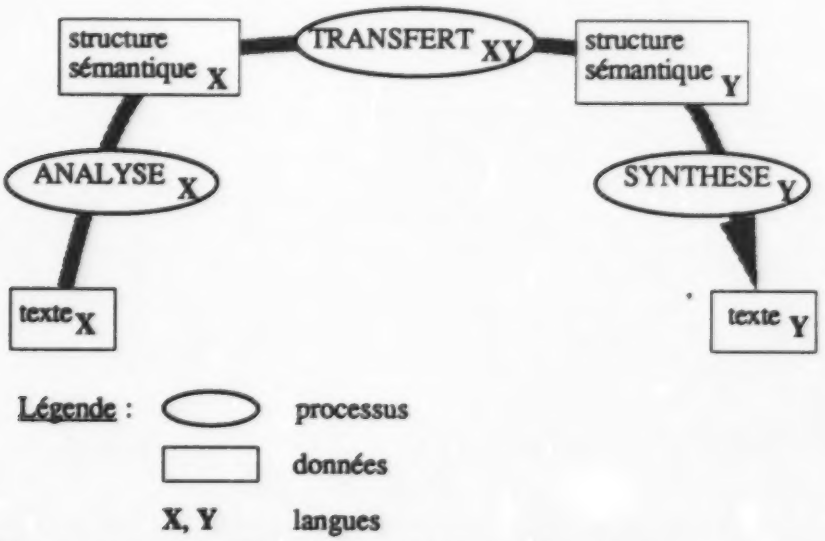


Figure 23 Décomposition schématique du processus de traduction.

Lorsqu'on exécute la commande de traduction (sélection de l'icone de la vache), dans le contexte d'un affichage des structures intermédiaires, le système ne réalise qu'une des trois

étapes, dépendant de l'état d'avancement déjà atteint. L'étape couramment exécutée est symbolisée par une inversion du contraste de l'icone rattaché à cette étape (cf. figure 24). Pour réaliser une traduction complète il est donc nécessaire, dans ce contexte, de sélectionner trois fois la commande de "traduction".

Pour avoir accès aux structures effectivement manipulées par le processus de traduction, il suffit de sélectionner l'icone correspondant à l'étape qui nous intéresse. Pour chacune des étapes, l'interface affiche un écran différent qui contient les données et les résultats relatifs à cette étape. C'est ce que nous allons illustrer à l'aide des figures 25 à 29 en suivant la traduction de la phrase `les cours de bovins de court engraissement baissaient de cinq dollars.`

Lorsqu'on affiche les informations relatives à chacune des étapes, en sélectionnant un des icones (cf. figure 24), on effectue une superposition d'un nouvel écran à l'écran principal d'IRMA. Cette superposition peut même atteindre trois couches comme nous le présentons dans les figures 28 et 29. Pour éliminer l'écran frontal et restituer le suivant dans l'ordre de superposition, il suffit de sélectionner "dans le vide" de l'écran à l'aide du bouton gauche de la souris.

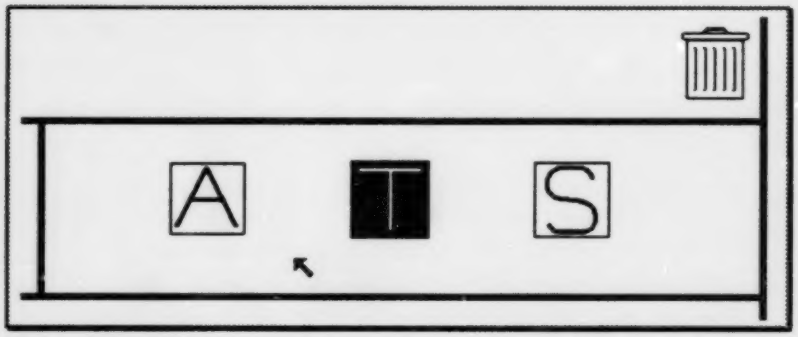


Figure 24 Configuration respective des icones lorsque le système effectue un transfert.

En sélectionnant l'icone A qui symbolise l'étape d'analyse (cf. figure 25), le texte à traduire est repris de l'éditeur vocal et apparaît dans la partie supérieure de l'écran. Le résultat est un graphe sémantique (cf. [1], pour des explications sur les structures sémantiques). À remarquer qu'une donnée apparaissant dans l'écran d'analyse et contenant des homophones ne subit pas l'effet secondaire mentionné dans *L'effet secondaire de la traduction*, section 2.3.

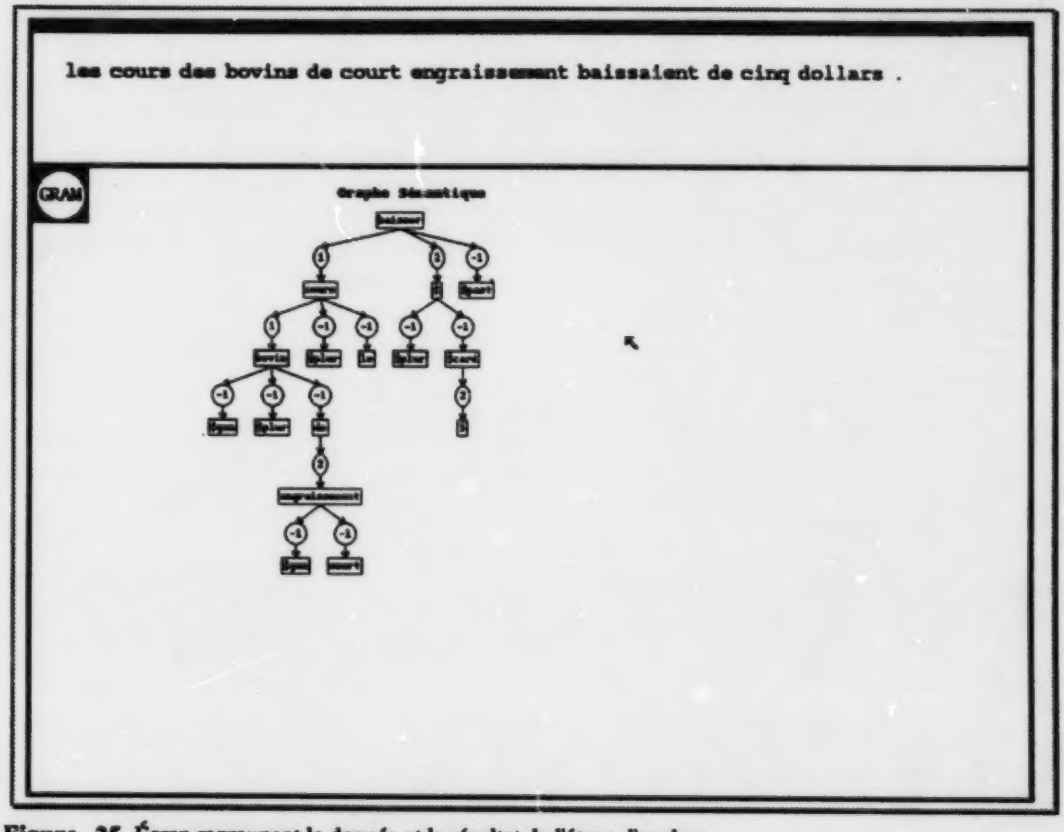


Figure 25 Écran regroupant la donnée et le résultat de l'étape d'analyse.

Le transfert, associé à l'icone `T`, (cf. figure 26) reprend comme donnée, dans la partie supérieure de l'écran, le résultat de l'analyse conformément au schéma de la figure 23. Seul l'espacement entre les noeuds est différent. Le résultat du transfert, dans la partie inférieure, est un autre graphe sémantique, cette fois, de la langue cible.

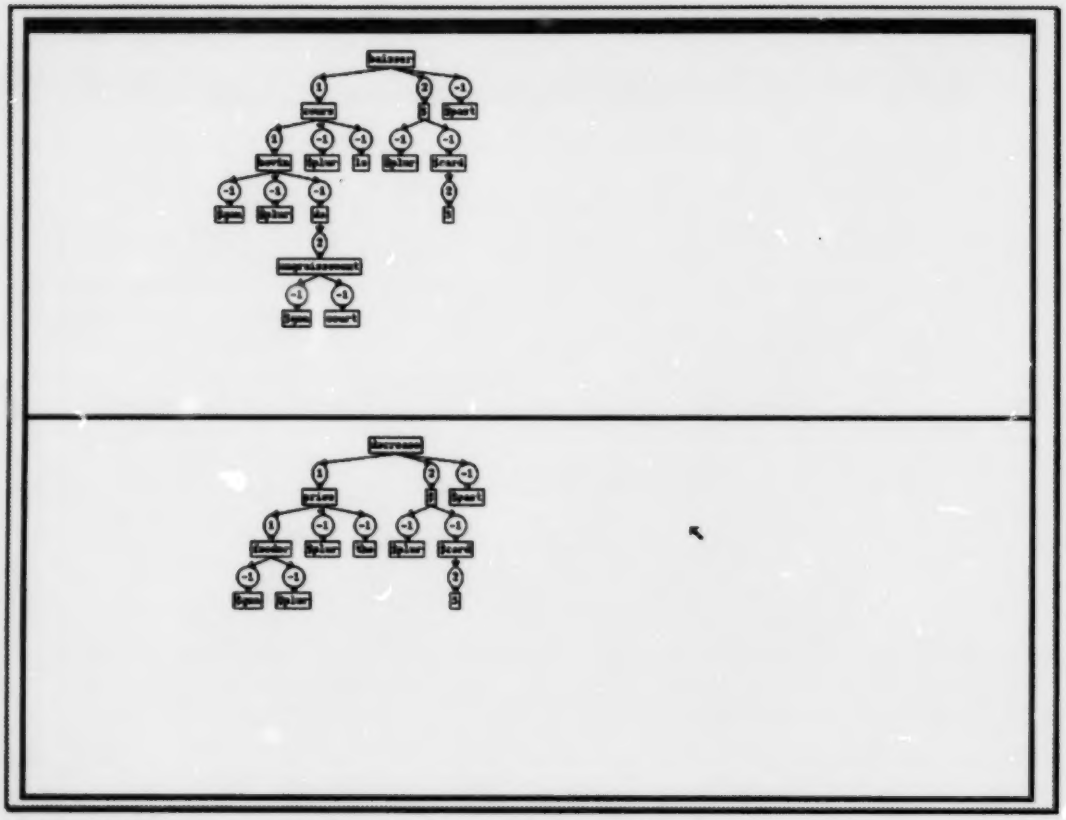


Figure 26 Écran regroupant la donnée et le résultat de l'étape de transfert.

La synthèse, associée à l'icone s, (cf. figure 27) reprend comme donnée, dans la partie supérieure de l'écran, le résultat du transfert conformément au schéma de la figure 23. Le résultat de la synthèse, dans la partie inférieure, est la phrase traduite. Cette phrase est en tous points identique à celle qui apparaît dans la fenêtre d'affichage des traductions (cf. section 2.4).

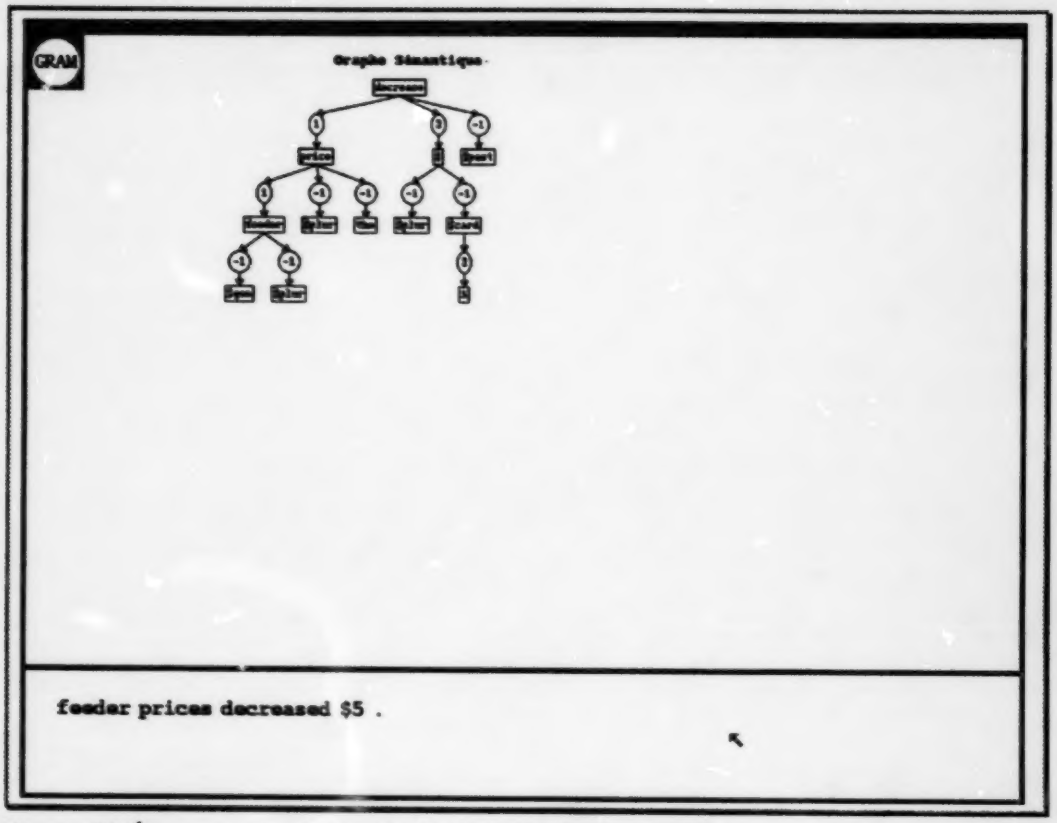


Figure 27 Écran regroupant la donnée et le résultat de l'étape de synthèse.

L'écran de l'étape d'analyse contient un icone en plus des données et résultats (cf. figure 25). Lorsqu'on sélectionne cet icone, on fait apparaître un nouvel écran dont la donnée est identique à celle du précédent, mais dont le résultat est une structure auxiliaire importante construite durant l'analyse: l'arbre syntaxique (cf. figure 28).

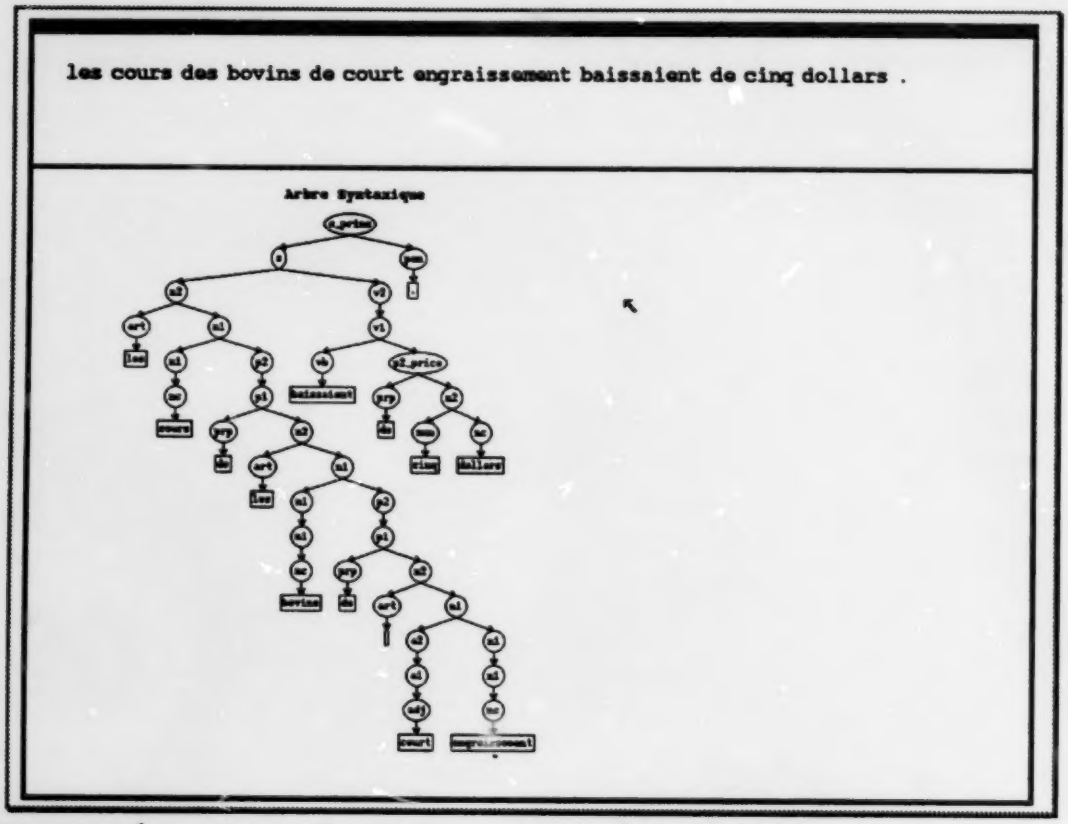


Figure 28 Écran auxiliaire contenant l'arbre syntaxique calculé lors de l'analyse.

L'écran de l'étape de synthèse contient aussi un icone en plus des données et résultats (cf. figure 26). Lorsqu'on sélectionne cet icone, on fait apparaître un nouvel écran qui contient l'arbre syntaxique dans sa partie supérieure et la phrase synthétisée dans la partie inférieure (cf. figure 29). La disposition des informations dans cet écran contrevient à la règle respectée jusqu'ici qui veut que la donnée d'une étape soit présente dans la partie supérieure et le résultat dans la partie inférieure. Le choix de cette disposition particulière reflète plutôt la relation de symétrie inverse qui existe entre l'analyse et la synthèse.

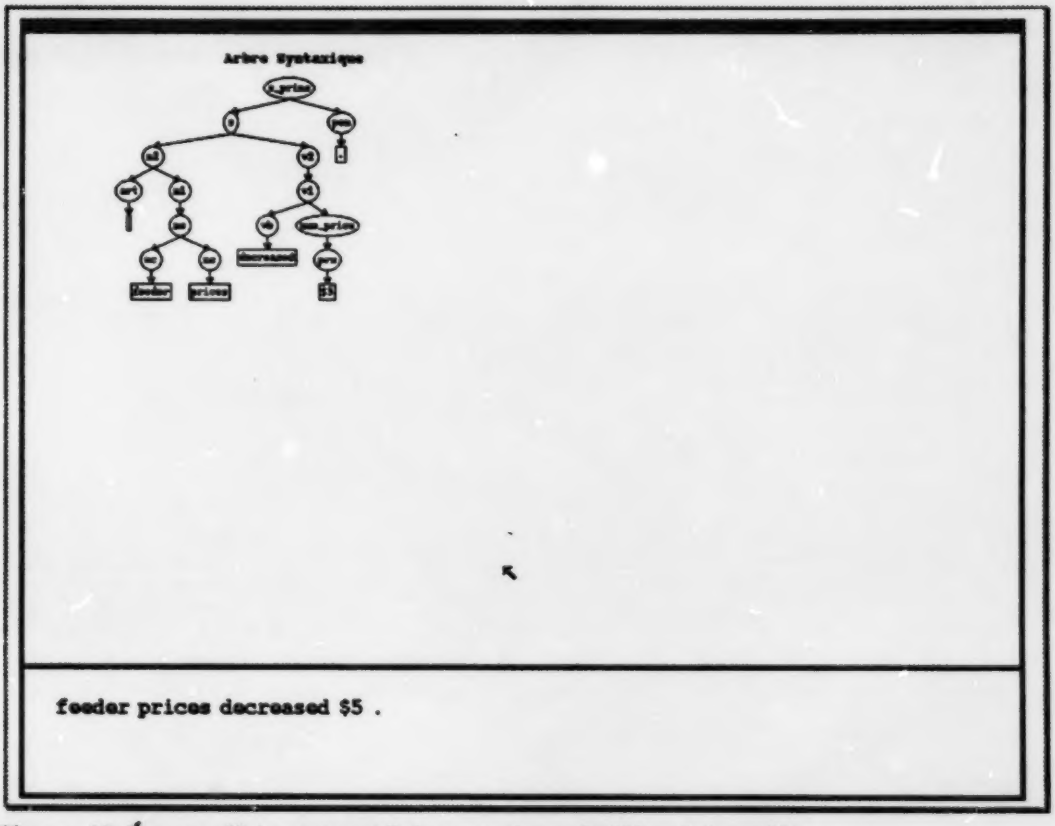


Figure 29 Écran auxiliaire contenant l'arbre syntaxique calculé lors de la synthèse.

Les idées directrices de la conception

3

Le développement d'IRMA s'est fait autour de quelques principes simples que nous tenons à présenter ici pour plusieurs raisons. Premièrement, un panorama d'IRMA n'est pas complet si nous omettons de parler des intentions originelles des réalisateurs de son interface d'utilisation. Deuxièmement, cette présentation nous force à faire une synthèse des idées et des principes que nous avons établis en matière de construction d'interfaces. Et troisièmement, nous pensons que cette présentation est une première étape vers l'établissement de principes généraux pour la construction d'interfaces *didactiques*.

3.1. Le profil des démonstrateurs et le public-cible

Nos premières préoccupations en matière de conception furent la caractérisation du profil des démonstrateurs et la détermination du public-cible.

Nous savions ainsi que les animateurs qui allaient se servir d'IRMA, étaient des étudiants universitaires, sans compétences particulières en informatique. Il ne s'agissait donc pas d'étudiants en informatique ou d'étudiants familiers avec ce domaine. Ainsi on ne pouvait espérer qu'ils s'adaptent facilement à un logiciel compliqué et qu'ils soient rapidement en mesure de faire des manipulations complexes. De plus, l'apprentissage des manipulations devaient se faire sur une seule journée, durant laquelle il fallait aussi les instruire sur la TAO, les thèmes à illustrer, les explications à donner. Ce qui est énorme et même probablement trop si l'on désire avoir de bonnes démonstrations. Ainsi, afin d'assurer le meilleur apprentissage possible, l'interface d'IRMA devait être conçue comme un outil simple et robuste et les manipulations devaient être intuitives et faciles à mémoriser.

À l'opposé des animateurs, les visiteurs d'une exposition comme Expotec forment généralement un groupe d'individus difficile à caractériser. On peut néanmoins dire qu'un tel groupe est de toute évidence moins homogène que celui d'une conférence internationale en linguistique informatique ou que celui d'une classe de secondaire. Ce qui revient à dire que ni les compétences dans un domaine spécialisé, ni l'âge des visiteurs ne sont des points communs au sein de ce groupe. Nous avons donc estimé que, si le public d'Expotec était aussi difficile à cerner, c'est que ses motivations, ses intérêts et ses attentes étaient multiples et divers. En conséquence, plutôt que d'illustrer uniquement un ou deux sujets simples, attrayants et percutants à l'aide d'IRMA, nous avons décidé de nous adresser à ces divers publics en survolant toute une variété de thèmes, allant de la synthèse vocale qui est probablement un des sujets les plus impressionnants, à l'explication des étapes de la traduction qui est peut-être le sujet le plus technique et le plus imperméable.

Nous pensons que cette approche permet à un public varié de trouver son compte dans une démonstration qui, à l'instar d'IRMA, englobe autant de technologies de pointe

diverses. Il était d'emblée certain que seule une minorité des spectateurs serait en mesure de comprendre tous les détails de ce que nous présentons à la console d'IRMA mais nous avons à tout prix refusé le principe du nivellement par le bas qui veut qu'en matière de vulgarisation, on s'adresse systématiquement au public le moins qualifié. Il n'en reste pas moins que le commentaire explicatif qui accompagne la démonstration est volontairement très intuitif et tente en général de faire appel à des notions connues ou du moins accessibles à la majorité (cf. annexe).

3.2. Les thèmes majeurs

Les difficultés de la TAO

La TAO est un domaine de technologie de pointe très complexe qui a probablement compté à l'heure actuelle plus d'échecs retentissants que de réussites spectaculaires. Une approche en laquelle nous croyons est la réalisation de systèmes spécialisés, donc aux compétences étroites, garantissant une traduction de haute qualité. CRITTER est de ceux-ci. Même si cette approche réductionniste augmente les chances de réussite, il existe une complexité inhérente au processus automatisé de la traduction que l'on ne peut pas éviter.

Nous avons décidé qu'une présentation de la TAO à travers IRMA devait donner l'occasion au spectateur de percevoir ou de se convaincre des difficultés de la tâche. La seule méthode éprouvée dont nous disposions pour atteindre ce but avant le développement d'IRMA, était l'affichage des structures intermédiaires qui sont transmises entre les différentes étapes du processus de traduction (cf. [3], pour les détails de la première mise en oeuvre de ces affichages). Nous avons donc décidé de reconduire cette méthode en l'adaptant à une nouvelle organisation de l'écran de la console d'IRMA.

Cette solution offre le double avantage suivant. Premièrement, dans la majorité des cas, il est fort probable que le spectateur ne soit pas en mesure de saisir tout ou partie de ce qui lui est présenté. Mais, comme la simplicité apparente des textes traduits par CRITTER laisse supposer que le problème est trivial, on peut au moins espérer que l'impression laissée par une présentation sur le fonctionnement du système sera celle correcte d'un problème intrinséquement complexe. Deuxièmement, pour quelques spectateurs avertis, cette illustration du fonctionnement de CRITTER est un moyen de parfaire leurs connaissances sur la TAO et de se convaincre si nécessaire qu'il n'existe pas de tricherie dans la démonstration.

La perte d'informations de la langue parlée par rapport à la langue écrite

Un aspect auquel tout locuteur que nous sommes est rarement sensible est la perte d'informations de la langue parlée par rapport à la langue écrite. Dans la mesure où IRMA fait usage d'un périphérique de reconnaissance vocale, il est tout à fait à propos d'illustrer ce phénomène, d'autant plus qu'il est simple à saisir. Cet aspect de la langue parlée est rattaché à l'existence des homophones hétérographes. À ce titre, le français illustre probablement mieux le problème que l'anglais, puisqu'en général le pluriel des noms réguliers et plusieurs terminaisons dans la conjugaison des verbes ne se prononcent pas.

La choix de faire apparaître cette particularité de la langue parlée répond à deux soucis constants dans notre cheminement. Premièrement, nous désirons faire preuve de rigueur scientifique: IRMA n'est pas une démonstration truquée dont le seul rôle est d'épater ou de mystifier le spectateur. Cette application doit refléter si possible l'état actuel des

technologies qui la composent en montrant leurs limites respectives. Deuxièmement, IRMA doit être une lentille grossissante sur des phénomènes de la langue naturelle, écrite ou parlée, reconnue ou synthétisée, que l'on ne perçoit qu'après y avoir été directement confronté.

La puissance de CRITTER en matière de désambiguïsation

Le développement de CRITTER n'a pas été motivé par la réalisation du système IRMA, mais IRMA est simplement un cadre possible d'utilisation de ce système de traduction. Ce cadre a l'avantage de permettre de tester CRITTER dans un environnement pour lequel il n'était pas prévu initialement: la traduction de textes entrés vocalement.

Un texte dicté se caractérise par une perte d'informations comparativement au texte équivalent écrit. Cette incertitude est due, comme nous venons de le dire, à l'existence d'homophones hétérographes. CRITTER a la capacité de résoudre la plupart de ces incertitudes grâce à son niveau d'expertise très élevé concernant la grammaire et la sémantique des textes concernés. L'aptitude de CRITTER à prendre en charge cette opération justifiait pleinement l'intégration d'un tel mécanisme dans IRMA, avec l'idée d'illustrer ainsi le potentiel de notre système de traduction.

3.3. L'organisation de l'interface

La mise en place de l'interface d'IRMA a nécessité une prise de décisions sur des problèmes ponctuels comme par exemple le positionnement de fenêtres à l'écran, le choix de jeux de caractères, le choix d'objets d'interaction, etc. Nous avons dû trouver des solutions particulières à chacun de ces problèmes, mais nous avons essayé tout au long de ce processus de garder à l'esprit quelques idées maîtresses autour desquelles nous avons voulu construire IRMA. Dans cette section, nous reprenons ces idées de base. Nous essayons d'argumenter un peu ces choix et nous décrivons par l'exemple ou en référence au chapitre précédent, quelques mises en oeuvre de ces principes.

Aucun des principes qui suit n'est fondé sur une étude approfondie de la littérature en ergonomie cognitive. Nous ne revendiquons que la mise en oeuvre de notre bon sens dans ce domaine, accompagné d'une expérience dans le développement de telles interfaces (cf. [3]).

Les principes de clarté et de sobriété

Une préoccupation constante lors du développement de l'interface d'IRMA fut de préférer les solutions *claires* et *sobres* aux solutions complexes et tape-à-l'oeil. C'est au cas par cas que ces principes ont été mis en valeur car il n'existe évidemment pas de méthode pour construire une solution claire et sobre. On peut seulement faire un choix parmi plusieurs solutions possibles en estimant qu'une des solutions est plus claire et plus sobre que les autres. Plutôt que d'essayer de définir précisément le sens que nous donnons à ces qualificatifs, nous allons illustrer par des exemples concrets ce que nous avons considéré comme clair et sobre.

Dans le cadre de l'éditeur vocal, nous avons volontairement et après mûres réflexions décidé de n'utiliser le logiciel de reconnaissance vocale que pour la saisie des mots qui composent la phrase. Un exemple bien connu d'utilisation d'un logiciel de reconnaissance vocale est la saisie vocale des commandes MS-DOS. La première solution pour l'éditeur

vocal envisageait que l'utilisateur puisse à la fois prononcer des mots de la phrase et des commandes d'édition du genre "début de phrase", "mot précédent", "effacer", etc. Ce mélange est tout à fait réalisable si on construit un mini-langage de commandes et il aurait été probablement facile de l'enseigner aux présentateurs. Par contre, pour le public, nous avons estimé que ce mélange de commandes et de données risquait de les embrouiller. Donc dans ce cas, nous avons préféré une implantation plus primitive mais sensiblement plus simple à assimiler.

De manière générale, la sobriété de l'interface d'utilisation d'IRMA est assez évidente si on compare le contenu des écrans d'IRMA à ceux de l'interface d'utilisation du système CRITTER [3]. Voici deux autres exemples pour lesquels le critère de sobriété a été décisif.

Premièrement, nous avons décidé qu'IRMA ne permettrait le traitement que d'une seule phrase à la fois. Nous avons écarté la possibilité de garder à l'écran plusieurs phrases et plusieurs traductions. La surcharge de textes à l'écran qu'aurait entraînée cette solution aurait considérablement nui à la lisibilité de l'écran et aurait forcément nécessité l'utilisation d'un jeu de caractères plus petit donc moins visible de loin. Par contre l'éditeur vocal que nous avons imaginé ne rappelle en rien un quelconque éditeur ou traitement de textes avec lequel un visiteur pourrait être familier.

Deuxièmement, toujours en comparaison avec l'interface de démonstration du système CRITTER, nous avons préféré ne pas mettre à disposition les commandes qui permettent d'intervenir dans le traitement effectué par Prolog. Ces commandes permettent de terminer inconditionnellement une exécution, de commander la recherche d'une autre solution et de resynthétiser une phrase lorsqu'on est en fin d'analyse et ainsi d'obtenir une paraphrase de la phrase originelle sans passer par une traduction. Ces commandes ont un sens lorsqu'on veut expliquer le détail des possibilités du sytème CRITTER. Il nous est apparu mal avisé et même néfaste de les introduire dans IRMA car il serait vain de vouloir atteindre un tel niveau de détails dans les explications si l'on ne s'adresse pas uniquement à un public de spécialistes.

Le langage des symboles graphiques

L'avènement des écrans graphiques et de la notion d'icone, de bouton, de sélecteur donne naissance à ce que l'on peut appeler un *langage de symboles graphiques*. La puissance de ce langage symbolique réside dans les références explicites qu'il crée entre le monde courant des objets qui sont suggérés graphiquement et le monde informatique des procédures activées lors de la sélection de ces symboles graphiques.

Nous avons décidé d'appliquer ce principe à IRMA dans l'espoir de rendre son abord plus immédiat et intuitif. Ainsi, dans le tableau de commandes, nous trouvons des sélecteurs qui permettent de sélectionner une valeur parmi une liste de valeurs disponibles, soit selon un mode circulaire, soit selon un mode d'accès direct. Dans ce cas, la référence à un bouton de syntonisation d'appareil de télévision est suffisamment claire pour que le spectateur puisse comprendre le sens de ces sélecteurs. Le tableau des paramètres de configuration contient des boutons en plus des sélecteurs. Là aussi le lien avec un bouton à presser est clair et il est aisé à l'utilisateur d'en saisir le mode d'utilisation.

Pour le reste de l'interface les commandes sont associées à des icones. La symbolique de l'icone réside uniquement dans l'image qu'elle contient et la référence qu'elle crée. Si une poubelle, un clavier et une fenêtre auxiliaire sont assez faciles et évidentes à représenter, une commande comme "traduire" est plus difficile à symboliser; il a donc fallu

innover. Comme dans l'interface de démonstration, c'est le sigle de CRITTER – "la vache" des rapports de marchés agricoles – qui joue ce rôle. Durant l'exécution d'une traduction, la vache est animée et cette animation reflète l'état du traitement informatique en cours. Le parallélisme entre les deux déroulements – la traduction et l'animation – crée une identification entre "traduire" et "manger" puisque "manger" est le thème de la petite animation. Ainsi si l'on veut une traduction, on va demander à la vache de manger en activant l'icone en question.

Le dernier point qui nous a posé un problème et pour lequel la solution est aussi graphique, était la représentation des homophones hétérographes dans l'éditeur vocal. Nous sommes parti du principe que le texte est *écrit* dans l'éditeur vocal et que ce qui est écrit est durable et vrai! Ainsi, selon cette logique, les homophones devaient apparaître sur quelque chose qui ressemble à un morceau de papier ou à un carton que l'on peut faire disparaître ou mieux déplacer. Ce qui est symboliquement caché sous le carton est la solution au problème élémentaire qui consiste à trouver la bonne écriture parmi les choix disponibles. Cette solution est calculée par CRITTER et elle apparaît lorsque l'interface simule le déplacement du carton.

Le flux des informations et la symétrie des affichages

La console d'IRMA sert principalement à montrer les données manipulées par le système. Une partie de ces informations est redondante: le texte reconnu en entrée vocale et le texte traduit et fourni à la synthèse vocale. Dans le cas de la copie écrite du texte dicté, il y a deux raisons à sa présence: l'écho nécessaire de la reconnaissance et l'illustration du problème des homophones. L'écho du texte dicté joue un rôle de modération – le présentateur attend l'écho d'un mot avant de prononcer le suivant – et d'homologation – l'affichage de la donnée écrite met en évidence ce que le système de reconnaissance a "compris". Dans le cas du texte traduit, il n'y a pas réellement de raison. Le seul argument que nous avons retenu est le besoin de symétrie dans l'affichage, qui se manifeste pratiquement par une organisation symétrique des informations à l'écran. L'écho de la reconnaissance (la donnée) est présenté dans la partie supérieure et le texte traduit (le résultat) est présenté dans la partie inférieure de l'écran. La distribution illustrée par cette première prise de contact avec IRMA définit très intuivement le flux des informations à travers l'écran, c'est-à-dire du haut vers le bas.

Dans les écrans plus techniques d'IRMA qui présentent les informations internes manipulées par le système, on a conservé ce même flux d'informations. D'autre part, la symétrie intrinsèque du processus de traduction, dans sa décomposition en étapes, nous a aussi permis de respecter une forme de symétrie sur les informations présentées dans ces différents écrans.

Pour augmenter la sensation de flux d'informations, pour chacune des étapes, il y a une distinction entre l'affichage de la donnée et du résultat. Lors de l'apparition de l'écran rattaché à l'étape en question, la donnée est présentée en bloc comme une information statique. Le résultat, par contre, est partiellement calculé à ce moment précis et est composé dynamiquement. Cette distinction est uniquement valable lorsqu'on fait apparaître ces informations pour la première fois. Si, en cours de présentation, on désire montrer les informations figurant dans un écran déjà présenté, les données et les résultats sont affichées en bloc, statiquement. Ce moyen permet de sensibiliser le spectateur au concept de *machine* qui part d'une donnée connue, qui calcule et qui affiche un résultat.

Les actions-réactions explicites

Un aspect toujours déroutant lors d'une démonstration, du point de vue du spectateur, est l'absence apparente d'une action engendrant une réaction visible; le spectateur ne voit que la réaction dont la cause lui échappe. Si une telle situation est isolée, le résultat n'est pas catastrophique, mais si elle se répète sur une séquence d'opérations, il est fort probable que le spectateur perde le fil de la démonstration.

Nous avons donc attaché beaucoup d'importance au respect de la règle qui veut que le spectateur puisse identifier facilement l'origine d'une réaction du système. Évidemment, la solution la moins coûteuse revient à exiger que le présentateur commente systématiquement les manipulations qu'il effectue. C'est probablement une excellente pratique mais nous avons voulu renforcer cet aspect en forçant le présentateur à effectuer des actions explicites.

La plus pure illustration de ce que nous appelons "action explicite" est la sélection à l'aide de la souris d'un icone dont le symbole présenté a une signification évidente. Le déplacement du curseur et l'arrêt sur l'icone sont des opérations visibles par le spectateur, et qui apparaissent comme des causes de ce qui suit dans la démonstration. Ce mécanisme d'interaction entre le démonstrateur et le système est évidemment plus lent et plus contraignant que la frappe d'une touche du clavier interprétée comme une commande. Voici quelques exemples d'actions-réactions explicites au sein d'IRMA.

La mise en service et hors service du clavier. La mise en service du clavier nécessite une opération de sélection d'un icone (cf. figure 17). Les réactions qui suivent immédiatement sont l'ouverture d'une petite fenêtre en lieu et place de l'icone et le déplacement automatique du curseur à l'intérieur de cette fenêtre afin de la rendre active. À ce moment, les caractères tapés au clavier sont affichés dans cette fenêtre. Pour mettre le clavier hors service, il suffit de déplacer le curseur hors des limites de la fenêtre ce qui est aussi une action visible.

La commande de traduction. Le fait d'associer la commande de traduction à un icone oblige un déplacement explicite du curseur dans l'espace de la fenêtre pour commander une traduction. Il aurait été possible de faire démarrer automatiquement la traduction en se basant sur la présence du point final de la phrase.

Malgré tout, IRMA n'est pas exempt de contre-exemples à cette règle et le développement de l'interface est passé par des étapes qui en contenaient d'autres qui ont heureusement été écartés. Voici un contre-exemple qui subsiste et un autre qui fait partie de ceux qui ont été rectifiés.

La fermeture des écrans attachés aux étapes de la traduction. À l'opposé de l'ouverture des écrans attachés aux étapes qui se fait selon notre principe en sélectionnant un icone, la fermeture se fait en "sélectionnant" un point quelconque en dehors de tout objet contenu dans la fenêtre. Il n'y a donc pas d'endroit précis attaché à cette commande. Cette opération peut surprendre le spectateur puisqu'il ne remarquera pas forcément la sélection du bouton gauche de la souris.

L'ouverture du tableau des paramètres de configuration. Dans un premier temps, c'est en sélectionnant le titre INTERFACE DE DÉMO IRMA que le présentateur faisait apparaître le tableau des paramètres de configuration. On voulait à ce moment cacher ces paramètres et éviter d'en parler durant la présentation. Hors, en introduisant le concept de système de reconnaissance vocale *mono-locuteur*, le nom du `présentateur` devient un paramètre d'exécution tout à fait justifié et sa sélection devient une des opérations du scénario de présentation (cf. annexe, acte 1, scène 1). Nous avons donc décidé d'associer un icone spécifique à l'opération d'ouverture du tableau.

L'utilisation simple de la souris

La souris des stations SUN est une souris à trois boutons. C'est probablement deux de trop si on se rapporte à ce qu'offre Macintosh. Même sans vouloir prôner l'utilisation de souris à un seul bouton, le contexte d'utilisation d'IRMA est tel que nous avons réduit l'utilisation de la souris SUN au seul usage du bouton de gauche. Le seul argument qui motive cette simplification est le profil des démonstrateurs établi plus haut.

3.4. La conception des outils

La centralisation des commandes des périphériques

La mise sur pied d'IRMA a nécessité l'utilisation d'une assez grande quantité de matériels différents (cf. chapitre 4). Il était donc impératif que cette profusion de matériels soit transparente à l'utilisateur d'IRMA et aux spectateurs. En ce qui concerne le spectateur, il fut décidé de cacher tout ce que l'on désirait dissimuler à l'intérieur du kiosque et de laisser croire que la console d'IRMA prenait tout en charge: la reconnaissance vocale, le traitement linguistique et la synthèse vocale. Pour simplifier la tâche de l'utilisateur et si l'on désirait que l'illusion soit vraiment parfaite, il fallait que la gestion des "périphériques" de reconnaissance et de synthèse vocales soit centralisée dans la console d'IRMA, au même titre que la gestion du traitement linguistique.

L'objectif principal de cette centralisation est essentiellement pratique. Le regroupement de toutes les commandes au sein d'un même écran simplifie évidemment l'utilisation et la gestion des périphériques vocaux. De plus, l'illusion créée par cette simplicité apparente de gestion des périphériques a un effet bénéfique sur la simplification de l'exposé. Aucun détail technique ou manipulation bizarre comme l'utilisation d'un autre clavier – celui du PC de reconnaissance (cf. section 4.2) – ne vient perturber la démonstration.

Le jeu restreint de commandes de l'éditeur vocal

L'éditeur vocal a constitué un problème majeur tout au long du développement de l'interface d'IRMA. La conception de cet éditeur a suscité beaucoup d'interrogations, en particulier sur le jeu de commandes que l'on mettait à disposition de l'utilisateur. Nous nous sommes finalement contentés d'un jeu minimal – celui décrit dans la section 2.3. Cette solution n'est pas immédiatement apparu comme une bonne, voire même peut-être la meilleure solution, en partie parce qu'en matière d'éditeurs les exemples qui nous entourent ne répondent pas aux critères que nous nous étions fixés, à savoir une grande simplicité d'utilisation et la possibilité de corriger facilement des erreurs de reconnaissance.

Voici un problème simple qui illustre bien la tendance automatique que l'on a de prendre exemple sur ce qui nous entoure: on s'est demandé si l'on allait introduire la notion de *curseur* qui se déplacerait dans le texte écrit. Il est clair qu'un curseur est, à priori, un concept que l'on retrouve dans tout éditeur. Dans notre cas, sa présence n'était pas complètement justifiée. Premièrement, l'utilisation prévisible de cet éditeur se limite à fournir un écho de la phrase dictée. Donc il suffit pour cette tâche d'avoir un curseur implicite qui reste toujours à la fin de la phrase en cours de création afin que le mot suivant soit automatiquement placé à la suite des autres. Deuxièmement, en cas d'erreur de reconnaissance, il est nécessaire de remplacer un mot. Donc, on a besoin de repérer un mot et non une position dans le mot, comme le permettrait un véritable curseur. La seule fonction que l'on s'interdit avec notre approche *sans curseur* est la possibilité d'insérer un mot entre deux autres mots de la phrase. Mais à nos yeux, cette fonction, somme toute très marginale dans la routine d'utilisation de l'éditeur vocal, ne justifiait pas tout l'investissement nécessaire à une telle mise en oeuvre.

Les détails techniques

4

Le kiosque IRMA qu'on a pu voir durant l'été 89 au Vieux-Port de Montréal renferme en fait beaucoup plus de matériel qu'il ne laisse paraître. En plus de la station de travail SUN 4/110, de son écran et de son clavier, du micro-casque et du haut-parleur, on peut apercevoir en ouvrant les panneaux arrières une unité de disque portative SUN, deux ordinateurs PC-compatibles Zenith 386 avec leurs claviers, un écran de PC, un synthétiseur DECtalk, un amplificateur audio, et deux blocs d'alimentation ("power bars") (cf. figure 1).

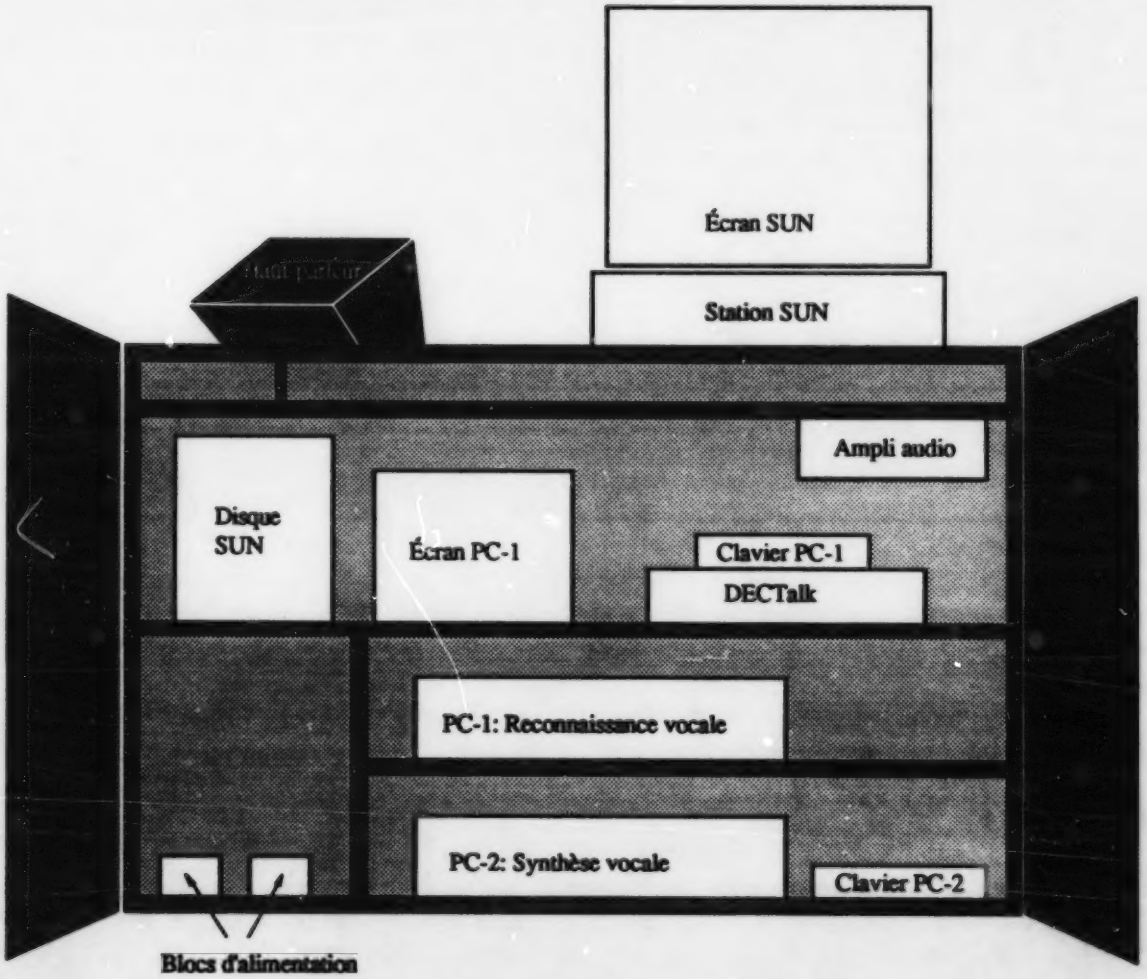


Figure 1 Schéma de la disposition du matériel à l'intérieur du kiosque IRMA.

Essentiellement, on peut diviser l'ensemble de ces équipements en trois groupes: *le groupe de reconnaissance vocale, le groupe de contrôle et de traitement linguistique* et *le groupe de synthèse vocale*. Le premier groupe comprend le micro-casque et un micro-ordinateur PC (le "PC de reconnaissance"), avec son clavier et son écran. Le deuxième groupe est constitué de tout l'équipement SUN (station de travail, clavier, souris, écran et disque portatif). Le troisième comprend le deuxième PC et son clavier (le "PC de synthèse"), l'appareil DECtalk, l'ampli audio et le haut-parleur.

La présence de trois ordinateurs indépendants dans la configuration d'IRMA (la SUN 4/110 et les deux PC) n'est pas le reflet d'un quelconque parallélisme dans le fonctionnement du système: les trois machines ne travaillent généralement pas toutes en même temps, et si elles le font, c'est uniquement le fruit d'un hasard qui, s'il est tout à notre avantage, n'en est pas moins indépendant de notre volonté. N'eût été des exigences en matériel de certaines des composantes que nous avions à notre disposition – les systèmes de reconnaissance vocale et de synthèse du français sont conçus pour un appareil de type PC, alors que le système CRITTER ne fonctionne que sur SUN –, c'est sûrement avec joie que nous aurions intégré toutes les opérations sous un même toit. En effet, la facilité d'utilisation, qui était une des exigences premières dans la conception d'IRMA, nous dictait de centraliser toutes les opérations de contrôle dans une même interface-usager. La répartition des tâches entre les trois machines nous a donc obligés à mettre sur pied des canaux et des protocoles de communication relativement complexes.

Le présent chapitre décrit la configuration interne de chacune des composantes des trois groupes, de même que leurs interrelations.

4.1. Le groupe de Contrôle et de Traitement

Le groupe de contrôle et de traitement assume principalement trois fonctions: la gestion de l'interface-usager, le traitement linguistique des données (en d'autres mots, la traduction des phrases), et la synchronisation des opérations avec les deux autres groupes de reconnaissance et de synthèse vocales.

Sur le plan matériel, la station SUN ne requiert pas d'équipement particulier (outre ses périphériques habituels), mais le logiciel nécessite beaucoup d'espace en mémoire (12 ou 16 Moctets de mémoire vive ne sont généralement pas superflus). De plus, un coprocesseur arithmétique ("Floating Point Unit") accélère grandement l'affichage de certaines constructions graphiques (l'installation d'Expotec 89 n'en possédait malheureusement pas).

Côté logiciel, le système CRITTER est entièrement écrit en Quintus Prolog [6], et l'interface-usager utilise l'environnement ProWINDOWS de Quintus [7], un paquetage ("package") de programmation orientée objet qui permet de développer des interfaces graphiques pour des programmes écrits en Quintus Prolog. Le tout fonctionne au sein de l'environnement SunView 1 [8], une interface multifenêtre du système d'exploitation SunOS [9].

Les activités des deux autres groupes sont gérées respectivement par le "PC de reconnaissance" et le "PC de synthèse". La station SUN est physiquement reliée à ceux-ci par des ports série, et synchronise les opérations au moyen de "messages". Bref, les deux PC peuvent être vus comme des périphériques de la station SUN dans la configuration d'IRMA. La nature des protocoles de communication employés est examinée plus à fond dans les sections qui suivent.

4.2. Le groupe de Reconnaissance

Lors de la dictée vocale d'une phrase à traduire, toutes les opérations de saisie et de reconnaissance des mots prononcés par l'usager sont effectuées par le PC de reconnaissance, qui est au centre de ce que nous appelons le groupe de reconnaissance. Au fur et à mesure qu'ils sont identifiés, les mots ainsi reconnus sont transmis au groupe de contrôle et de traitement linguistique, qui gère l'éditeur vocal (cf. section 2.3). À son tour, le groupe de contrôle peut envoyer des messages au groupe de reconnaissance, dans le but de changer le mode d'opération de ce dernier.

Le PC de reconnaissance a donc deux tâches à accomplir simultanément: gérer les opérations de saisie et de reconnaissance vocale en tant que telles, et assurer la communication bidirectionelle avec le groupe de contrôle. La première de ces tâches est effectuée par le système VoiceScribe-1000, que nous décrivons dans les deux sections qui suivent. La deuxième est prise en charge par un ensemble de programmes que nous décrivons dans la troisième section.

Le système VoiceScribe-1000 et le logiciel DragonKEY

VoiceScribe-1000[1] est un système mono-locuteur de reconnaissance vocale de mots isolés. Il s'agit d'un système commercial qui se présente sous la forme d'une carte à insérer dans un PC, d'un micro-casque que l'on connecte directement à la carte, et d'une librairie de logiciels servant à exploiter les possibilités de la carte.

La carte VoiceScribe-1000 transforme le signal électrique du microphone en un signal digital, et toutes les opérations de reconnaissance sont effectuées par la partie logiciel du système. Nous avons utilisé à cet effet le logiciel DragonKEY [10], fourni avec la carte, dont la fonction première est de permettre le contrôle d'applications PC à partir d'un microphone.

DragonKEY effectue grosso modo la reconnaissance vocale en deux étapes, qui correspondent à la consultation de deux tables distinctes. Dans un premier temps, le système tente d'associer le signal digital reçu de la carte à une étiquette représentant un mot connu du système. Ceci correspond à une recherche dans une table de couples signal-digital/étiquette, table qui est appelée le "vocabulaire" de l'application. Dans un deuxième temps, on associe à l'étiquette trouvée une chaîne de caractères, laquelle est utilisée comme valeur de sortie de DragonKEY. Cette deuxième étape est le résultat d'une recherche au "lexique", qui fait lui-même partie de la "grammaire" de l'application. Cette grammaire (régulière) définit l'ensemble des suites d'étiquettes valides, et permet de réduire l'espace de recherche lors de la reconnaissance.

C'est le gestionnaire du système qui a la responsabilité de fournir à DragonKEY la grammaire (incluant le lexique) de son application. Quant au vocabulaire, il est fourni par l'usager lorsque celui-ci "entraîne" DragonKEY à "reconnaître sa voix": il s'agit de répéter 4 ou 5 fois devant le micro la "chaîne vocale" correspondant à chaque étiquette du lexique. Comme le système est mono-locuteur, chaque usager doit constituer son propre vocabulaire.

---

1 VoiceScribe-1000 est une marque déposée de Dragon Systems, Inc.

À l'utilisation, DragonKEY se présente comme un programme résident, c'est-à-dire qu'il reste inactif en mémoire pendant l'exécution d'autres applications, en attente d'un signal prédéterminé. Ce signal est caractérisé par l'arrivée d'information en provenance de la carte VoiceScribe. Le programme met alors en marche les mécanismes de reconnaissance, et écrit sur le tampon du clavier ("keyboard buffer") la chaîne de caractères associée au signal digital reconnu. Ainsi, la reconnaissance d'un mot par DragonKEY produit exactement le même résultat que si l'usager avait tapé les caractères de la chaîne correspondante au clavier.

L'environnement VoiceScribe pour IRMA

Tel que mentionné plus haut, IRMA utilise le système VoiceScribe-1000 avec le logiciel d'exploitation DragonKEY pour la dictée vocale des phrases qu'IRMA doit traduire. Lorsque l'usager allume le PC, le fichier d'initialisation ("autoexec.bat") s'occupe d'activer la carte Dragon et de lancer le programme résident, de même que les programmes responsables de la communication entre le PC de reconnaissance et le reste de la machinerie d'IRMA. Nous discuterons de ces programmes dans la prochaine section. Voyons maintenant de quoi DragonKEY a besoin pour effectuer correctement la reconnaissance des phrases dictées par l'usager.

Puisqu'IRMA traduit aussi bien de l'anglais vers le français qu'inversement, la reconnaissance doit être possible dans les deux langues. Ainsi, on doit fournir deux grammaires distinctes au système: une pour l'anglais et une pour le français. Le système CRITTER connaît présentement plus de 2 000 formes de mots en anglais, et plus de 5 000 formes en français. Puisque d'une part le système VoiceScribe-1000 ne peut gérer un vocabulaire de plus de 1 000 mots (d'où son nom), et que d'autre part, IRMA n'existe que pour des fins de démonstration, nous avons décidé de limiter la taille des lexiques à une fraction des capacités de CRITTER: on reconnaît dans chaque langue un peu plus d'une centaine de formes.

Notons que c'est à ce niveau que sont identifiés les homophones hétérographes (cf. section 2.3). Du point de vue de la description grammaticale, un ensemble d'homophones hétérographes se rattache à une seule étiquette, mais la chaîne de caractères qui lui est associée représente la liste de toutes les graphies possibles. C'est donc au gestionnaire du système de reconnaissance que revient la tâche d'identifier les mots qui ont des prononciations similaires ou suffisamment semblables pour être systématiquement confondues.

Dans le cadre de la dictée vocale pour IRMA, nous avons opté pour une grammaire que l'on pourrait qualifier de "triviale": toute suite de mots du lexique constitue une phrase acceptable. En terme d'expressions régulières, `<phrase> = <mot>*`. Ceci implique que la validation grammaticale des phrases (validation de la position relative des mots à l'intérieur d'une phrase dictée) n'est pas faite au moment de la saisie. On laissera le système CRITTER s'occuper de cette tâche.

Bien sûr, cette façon de procéder n'a rien de particulièrement avantageux pour l'exercice de reconnaissance, puisqu'à tout moment, tous les mots du lexique sont "actifs", i.e. qu'ils sont tous des candidats potentiels. Toutefois, deux facteurs justifient une telle approche. Premièrement, les fonctions de l'éditeur vocal (cf. section 2.3) supposent que l'on peut à tout moment revenir à n'importe quel point de la phrase dictée et redicter un mot. Pour que cette "redictée" soit possible, il faut pouvoir resituer l'analyseur vocal dans le bon contexte

grammatical, chose compliquée vu les possibilités de DragonKEY (à moins bien sûr qu'il n'y ait qu'un seul contexte grammatical possible, ce qui est le cas avec notre grammaire triviale).

Deuxièmement, le système de traduction automatique CRITTER auquel sont ultimement envoyées les phrases dictées, a à sa disposition des grammaires du français et de l'anglais extrêmement sophistiquées. D'abord, l'idée de devoir reproduire cette information au niveau de l'analyseur vocal nous paraît particulièrement inélégante et lourde: toute modification aux grammaires de CRITTER (elles sont fréquentes) entraîne automatiquement des modifications équivalentes dans les grammaires vocales. De plus, le langage de description syntaxique de VoiceScribe est loin d'avoir la puissance expressive des formalismes grammaticaux de CRITTER. Au mieux, nous pouvons espérer construire des grammaires VoiceScribe définissant des langages qui incluent ceux définis par les grammaires correspondantes de CRITTER. Mais cette entreprise, dans la mesure où elle tente d'aller plus loin qu'une grammaire triviale, s'avère rapidement d'une complexité rébarbative.

De toute évidence, la solution idéale réside dans l'utilisation directe des grammaires de CRITTER pour contraindre les recherches de la reconnaissance vocale. Hélas, l'analyse grammaticale de CRITTER se fait de façon descendante ("top-down"), et par retour arrière ("backtracking"). Or, cette méthode se prête très mal à une validation en temps réel, telle que celle qu'exige normalement la dictée vocale.

En plus des grammaires française et anglaise, le système de reconnaissance vocale doit avoir à sa disposition les vocabulaires de tous les usagers potentiels du système, et ce pour chacune des deux langues. Ces vocabulaires sont créés par les usagers eux-mêmes, à partir de la "console" de DragonKEY. Il est intéressant de noter que l'usager n'est pas obligé d'entraîner le système à reconnaître sa voix pour la totalité des mots des deux lexiques. S'il compte n'utiliser les capacités de dictée vocale d'IRMA que pour quelques phrases choisies, il peut se contenter de n'entraîner VoiceScribe que pour les mots qui apparaissent dans ces phrases.

Les communications entre les groupes de Reconnaissance et de Contrôle

Du point de vue du PC de reconnaissance, les communications avec le groupe de contrôle peuvent se résumer ainsi: d'une part, il reçoit sur le port série des messages du groupe de contrôle, qui sont soit des messages de commande, soit des messages de synchronisation. Les messages de commande s'adressent généralement au programme DragonKEY: changement de grammaire ou de vocabulaire, mise en-circuit ou hors-circuit du microphone, etc. Les messages de synchronisation s'adressent plutôt au dispositif de communication lui-même: demandes d'accusé de réception ("acknowledgement"; voir plus loin) et autres signaux. D'autre part, il envoie des messages au groupe de contrôle, qui peuvent être soit des mots reconnus par le système VoiceScribe, soit des signaux divers, tels des accusés de réception, des avertissements signalant l'échec d'une commande envoyée à DragonKEY, etc.

Le programme de communication comme tel est constitué de deux processus: un "récepteur" et un "émetteur". Ces deux processus fonctionnent en parallèle au sein de l'environnement MultiC [5], qui effectue une gestion multi-tâches de programmes C. L'émetteur lit sans arrêt le contenu du tampon du clavier, et reproduit à peu près

intégralement son contenu sur le port série, à l'intention du groupe de contrôle. C'est ainsi que les mots reconnus par DragonKEY sont transmis à l'éditeur vocal.

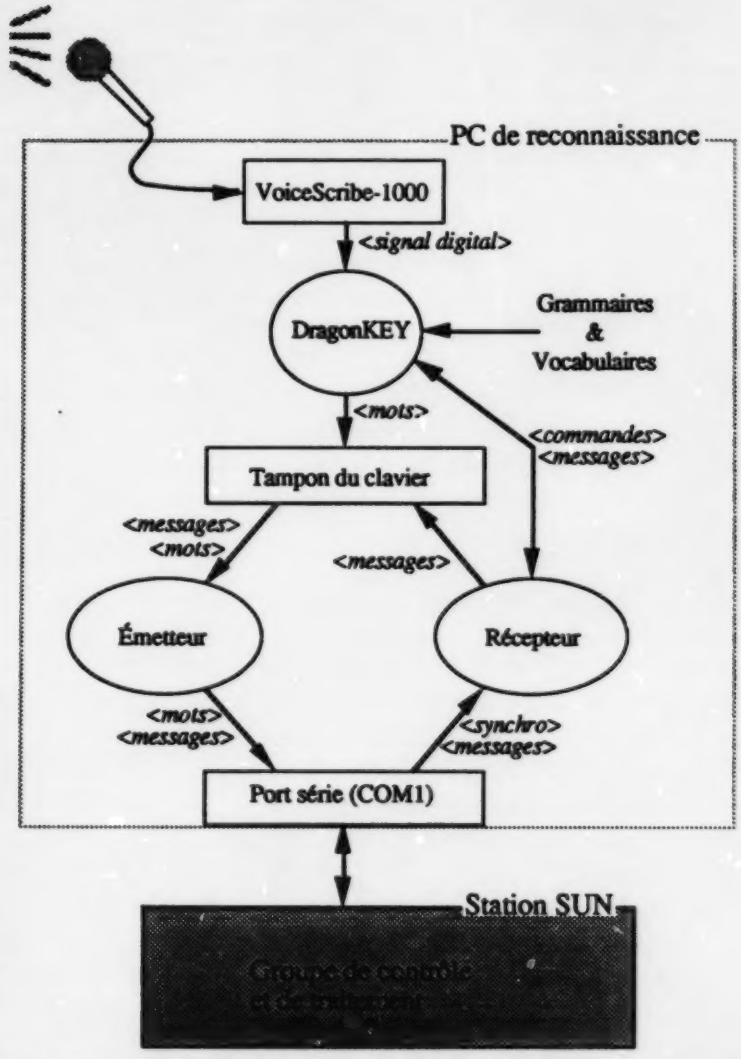


Figure 2 Schéma du flux des données entre les groupes de Reconnaissance et de Contrôle.

Le récepteur, quant à lui, lit sans arrêt le contenu du port série, et en trie le contenu. Les messages de synchronisation sont généralement traités à même le récepteur. Prenons par exemple la demande d'accusé de réception mentionnée plus haut: il s'agit d'une façon pour le groupe de contrôle de vérifier que le groupe de reconnaissance est bien à l'écoute, et en mesure de fournir de l'information. Lorsque le récepteur reçoit une telle demande, il envoie un message sur le tampon du clavier. Celui-ci est ensuite lu par l'émetteur qui le retransmet automatiquement au groupe de contrôle. On a ainsi bel et bien vérifié que la boucle est active. Le bouton **TEST** de l'interface-usager (cf. *Les paramètres de configuration*, dans la section 2.2) déclenche la demande d'un tel accusé de réception. Le test réussit si cet accusé de réception arrive au groupe de contrôle à l'intérieur d'un intervalle de quelques secondes.

Notons d'autre part que chaque fois que le groupe de contrôle envoie un message au groupe de reconnaissance, ce message est suivi d'une demande d'accusé de réception. C'est une façon simple de vérifier que tout s'est bien passé.

Lorsque le récepteur reçoit un message de commande pour DragonKEY, il lance un signal d'interruption à l'intention de ce dernier, et transmet la commande en l'écrivant à un endroit précis de la mémoire (c'est le protocole selon lequel DragonKEY reçoit ses commandes). Il attend ensuite le résultat de cette commande (succès ou échec), et en cas d'échec peut faire suivre le message d'erreur au groupe de contrôle en l'écrivant sur le tampon du clavier (cf. figure 2).

La liste qui suit énumère l'ensemble des types de messages transitant par le câble série reliant les deux groupes. Le protocole est relativement simple: le premier caractère du message identifie son type, et un séquence de caractère prédéfinie marque la fin. Tout ce qui apparaît entre les deux (_ dans la liste) constitue le corps du message. Deux types de message n'ont pas besoin de marqueur de fin, soient les <^A> et <^C> émis par le groupe de contrôle: ce serait inutile, puisqu'ils sont constitués chacun d'un seul caractère, ayant une signification spéciale pour le récepteur du groupe de reconnaissance.

groupe de Reconnaissance → groupe de Contrôle:

- `#_<LF>` chaîne de caractères associée à un mot reconnu par VoiceScribe,
- `<^B><LF>` accusé de réception,
- `<^B>_<LF>` message d'ordre général (avertissement d'échec de commande, etc).

groupe de Contrôle → groupe de Reconnaissance

- `#_<NUL><LF>` commande à DragonKEY avec les valeurs possibles suivantes pour _:
  - `console` démarrage du programme de "console" de DragonKEY permettant (entre autres) la constitution des tables de vocabulaire par les usagers,
  - `microphone on/off` mise en-circuit ou hors-circuit du microphone,
  - `overlay french/english` changement de langue source (spécification de la grammaire VoiceScribe à utiliser),
  - `vocabulary _` identification de l'usager (spécification du vocabulaire VoiceScribe à utiliser);
- `<^A>` demande d'accusé de réception,
- `<^C>` demande de re-synchronisation (le récepteur du groupe de reconnaissance oublie tout caractère reçu à ce point et attend un nouveau message).

4.3. Le groupe de Synthèse

Les phrases traduites par CRITTER peuvent être reproduites sous forme vocale par le groupe de synthèse d'IRMA. Celui-ci utilise à cet effet deux systèmes distincts de synthèse vocale, soient les systèmes DECtalk (pour la synthèse de l'anglais) et TELEVOX (pour le

français). Cette dichotomie illustre bien l'état de l'art dans le domaine de la synthèse: les systèmes qui sont présentement disponibles produisent des résultats parfois étonnants, mais sont souvent conçus spécifiquement pour une langue donnée, et ne peuvent être reprogrammés.

DECtalk

L'appareil DECtalk[1] se présente comme un périphérique (au même titre qu'une imprimante), qui accepte en entrée des "phrases" anglaises, et les reproduit en sortie sous une forme "parlée". L'entrée se fait par le biais d'un lien série, typiquement avec un ordinateur, et la sortie soit par un haut-parleur intégré à l'unité, soit par une sortie audio standard (du type que l'on branche dans un appareil haute-fidélité), soit par une sortie de type téléphonique.

Quoique limité à la synthèse de l'anglais, DECtalk offre à l'usager une grande flexibilité quant au réglage de la voix – voix féminines ou masculines, débit d'élocution, expressivité, etc. – et à la finesse de prononciation. IRMA utilise l'appareil de la façon la plus simple: on se sert de la voix de défaut ("Perfect Paul"), et on se contente d'envoyer directement les phrases à dire, sans aucun souci d'améliorer leur prononciation.

TELEVOX

Le système TELEVOX[2] de synthèse vocale du français se présente à peu près comme le système de reconnaissance VoiceScribe (cf. section 4.2): il s'agit d'une carte complémentaire de format PC, accompagnée d'un logiciel d'exploitation. Comme VoiceScribe, le tout s'intègre à MSDOS au moyen d'un driver. Lors de la synthèse vocale d'une phrase, le programme de communication, décrit brièvement dans la section suivante, appelle le driver et lui fournit la phrase à reproduire. Le logiciel de TELEVOX analyse celle-ci et en déduit une série de "trames", i.e. de paramètres utilisables par la carte de synthèse. La carte transforme ces trames en un signal analogique prêt à être amplifié.

Les phrases sont fournies au système sous la forme de chaînes de caractères ASCII normales. L'usager a la possibilité d'insérer à l'intérieur de celles-ci certains "caractères de contrôle" ayant pour but de modifier la prosodie de la phrase. IRMA ne tient pas compte de ces possibilités et transmet des phrases littérales à la synthèse.

Les communications à l'intérieur du groupe de Synthèse et avec le groupe de Contrôle

Le contrôle des opérations de synthèse est assuré par un PC (le "PC de synthèse"). D'une part, celui-ci gère la répartition du travail et la communication avec les deux systèmes de synthèse. D'autre part, il prend à charge la communication avec le groupe de contrôle et de traitement linguistique.

Les communications entre le groupe de contrôle et le groupe de synthèse sont supportées par un câble série, qui lie la SUN au PC de synthèse. Elles sont unidirectionnelles, et composées d'un seul type de messages, à savoir: des phrases à être synthétisées par DECtalk ou TELEVOX. Chaque phrase est précédée d'un caractère de

---

1 DECtalk est une marque déposée de DIGITAL Equipment Corporation, Inc.
2 TELEVOX est distribué par ELAN Informatique, sous licence du CNET.

contrôle qui indique s'il s'agit d'une phrase en français (caractère £) ou en anglais (caractère a), et se termine par un code de fin de phrase (le caractère ~). Notons qu'à l'intérieur de ces phrases, les caractères accentués sont encodés selon la convention ASCII étendue standard d'IBM. La communication est gérée par un programme en C qui lit en permanence le port série du PC, et achemine les messages reçus vers l'un ou l'autre des deux systèmes de synthèse.

L'appareil DECtalk est connecté au PC de synthèse par un lien série. Si la phrase à synthétiser est en anglais, elle est donc acheminée au deuxième port série du PC (le premier étant utilisé pour le lien avec le groupe de contrôle). Quant aux phrases françaises, elles sont soumises à la carte TELEVOX de façon interne (cf. figure 3).

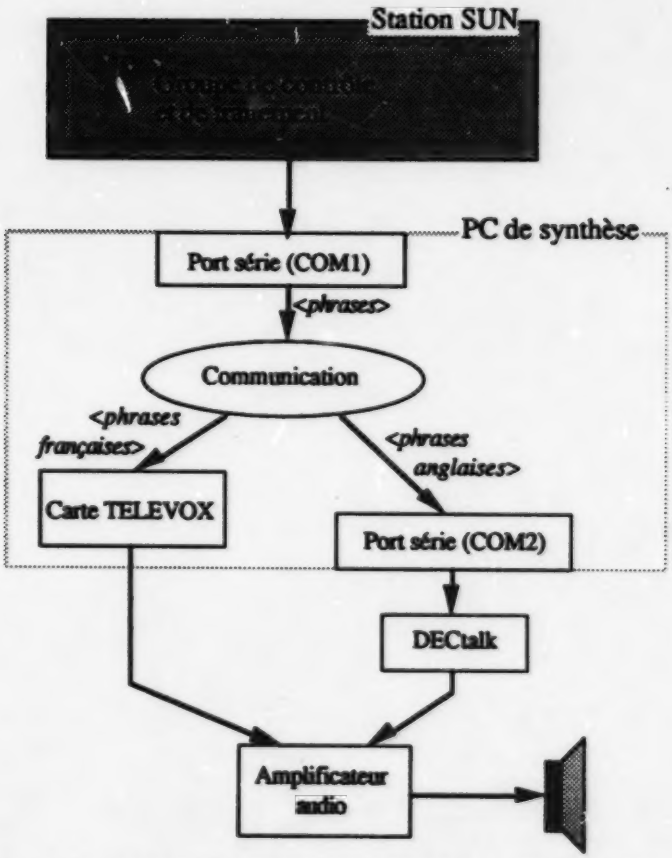


Figure 3 Schéma du flux des données entre les groupes de synthèse vocale et de contrôle.

Tel que mentionné plus haut, l'un et l'autre systèmes de synthèse vocale permettent à l'usager de régler certains de leurs paramètres opérationnels. Ces paramètres sont ajustés par le programme de communication du groupe de synthèse lors de la mise en circuit du PC de synthèse, et ne peuvent être modifiés à partir de l'interface IRMA.

Les sorties audio de DECtalk et de TELEVOX sont connectés à l'ampli audio, auquel est intégré un mélangeur trois voix (la troisième voix est utilisée pour amplifier les éventuels "bips" du groupe de reconnaissance). On peut donc calibrer séparément les niveaux sonores de sortie des deux systèmes de synthèse.

Les variantes d'IRMA

5

Même après son installation au Vieux-Port de Montréal pour Expotec, IRMA a subi des modifications et a été l'objet d'expériences. Nous décrivons ici les deux plus intéressantes.

5.1. L'expérience Tubach

Tel que mentionné à la section 4.2, la reconnaissance vocale d'IRMA repose sur l'utilisation de la carte VoiceScribe-1000. La dictée vocale à l'aide de cette carte est plutôt pénible, ceci pour deux raisons: premièrement, l'élocution par mots isolés n'a rien de particulièrement agréable; deuxièmement, les erreurs de reconnaissance sont très fréquentes (il n'est pas rare d'effectuer 4 ou 5 répétitions à l'intérieur d'une phrase d'une dizaine de mots). Ces erreurs peuvent être attribuées à l'absence de contraintes syntaxiques dans l'exercice de la reconnaissance – utilisation d'une grammaire triviale – (cf. *L'environnement VoiceScribe pour IRMA*, dans la section 4.2) et au niveau de bruit ambiant très élevé durant les démonstrations, mais la faute revient aussi au fait que le système VoiceScribe est bien loin de représenter l'état de l'art en matière de reconnaissance.

La visite de Jean-Pierre Tubach, professeur à Télécom Paris (École Nationale des Télécommunications – ENST) au CCRIT durant l'été 89 a donné lieu à une expérience intéressante: celui-ci a profité de son séjour sabbatique parmi nous pour tenter de fournir à IRMA une interface d'entrée vocale plus avancée que celle qu'elle possédait déjà. Deux systèmes ont été testés à cet effet, soient les systèmes VECSYS Datavox et X-COM Media50, qui se présentent tous deux comme des systèmes "complémentaires", à installer dans un micro-ordinateur de type PC. Le premier est un système mono-locuteur de reconnaissance en parole continue. Le second fonctionne aussi en parole continue, mais est multi-locuteur. Les détails de ce travail apparaissent dans [4].

Les modifications faites à l'interface-usager d'IRMA ont été réduites au minimum pour l'expérience, de telle sorte que celle-ci ne se distingue d'aucune façon de l'interface originale. La seule différence sensible dans cette variante d'IRMA est le mode de dictée: les phrases à traduire sont ici dictées par "membres de phrases". Ainsi, on peut énoncer au microphone la phrase suivante (/ désigne une pause facultative dans l'élocution, // une pause obligatoire):

```
Les prix de gros / des vaches de boucherie // ont augmenté / de
trois dollars le kilo // à Montréal // en raison d'arrivages très
faibles // point final
```

En fait, la dictée par phrases complètes aurait été possible avec l'un et l'autre des systèmes de reconnaissance testés, mais il fût déterminé que le mode par membres de phrases était le plus approprié pour ce genre d'application.

Les résultats de la reconnaissance, pour un système comme pour l'autre, sont pour le moins étonnants, et de loin supérieurs à tout ce qui pouvait être espéré avec un système du

genre de VoiceScribe-1000. Et pour cause: il s'agit d'une toute autre technologie. Toutefois, il faut dire à la décharge de VoiceScribe que le succès de l'expérience de la reconnaissance en parole continue repose aussi sur l'utilisation de grammaires très restrictives, solution que nous avions écartée dans la mise sur pied du groupe de reconnaissance. Ainsi, dans cette version d'IRMA, la dictée vocale est non seulement limitée à une fraction du vocabulaire connu de CRITTER (ce qui était aussi le cas pour VoiceScribe), mais aussi à une fraction des constructions syntaxiques admises par l'analyseur du système de traduction.

L'expérience démontre néanmoins que la dictée vocale pour un environnement tel que celui d'IRMA peut être plus qu'un "jouet". Dans la perspective d'un poste de travail pour traducteur basé sur la technologie CRITTER, les résultats obtenus laissent envisager éventuellement la possibilité pour le traducteur de dicter des traductions ou des corrections comme alternative ou comme complément à la traduction automatique.

5.2. IRMA Multi-processus.

Le programme Prolog chargé de la gestion de l'interface-usager, du traitement linguistique des données et de la synchronisation des opérations de saisie et de synthèse vocales est conçu comme un unique processus. Ainsi, lorsque l'usager commande une traduction en sélectionnant l'icone de la vache, le programme lance l'exécution de prédicats des modules linguistiques du système, et la gestion de l'interface est interrompue jusqu'à la fin de l'opération commandée. Si l'usager effectue d'autres manipulations à l'écran durant cet intervalle, celles-ci sont mémorisées et effectuées dans l'ordre lorsque le programme reprend le contrôle de l'interface.

Une des conséquences de cette architecture est l'impossibilité pour l'usager d'interrompre une traduction, autrement qu'en suspendant l'exécution du programme en entier. Ce type d'interruption est pourtant très utile, puisqu'il n'est pas rare que le système "s'emballe" lors du traitement de phrases agrammaticales ou particulièrement complexes. Au lieu des quelques secondes habituelles, une traduction peut alors nécessiter plusieurs minutes de traitement et, dans le meilleur des cas, se solde par un échec. Inutile de mentionner qu'au cours d'une démonstration, un tel phénomène ne manque pas d'être extrêmement embarassant!

Une des solutions à ce problème est le fractionnement du programme principal en deux ou plusieurs programmes qui sont exécutés comme des processus distincts. Cette approche est détaillée dans [3].

La réalisation d'une version d'IRMA "multi-processus" a été entamée, dans laquelle la partie gestion de l'interface du système constitue le processus "maître", et la partie traitement linguistique est exécutée en "esclave". Pour des raisons que nous exposons plus loin en détail (cf. chapitre 6), la mise en oeuvre de cette version n'a pas été achevée. La séparation des tâches en deux programmes et la mise sur pied de protocoles de communication inter-processus ont été réalisés, mais les travaux ont été interrompus au moment de modifier l'apparence de l'interface-usager pour réfléter les nouvelles possibilités du système (travail estimé à une demi-journée). Toutefois, les avantages d'une telle implantation sont assez clairs.

Le premier de ces avantages est la gestion ininterrompue de l'interface pendant l'exécution des traductions. Lorsque l'usager demande à IRMA d'effectuer une traduction,

le processus de gestion d'interface délègue la requête au processus de traitement linguistique, et continue de gérer l'interface en attendant que l'esclave ait fini le travail. L'usager est donc libre d'effectuer d'autres manipulations pendant ce temps, telles que la dictée d'une autre phrase, une demande de changement de langue-source, etc. Lorsque l'esclave a terminé le travail demandé, le maître récupère l'information et l'affiche.

Deuxièmement, si la machine sur laquelle tourne le processus maître fait partie d'un réseau de stations SUN, il est possible de lancer l'exécution de l'esclave sur une autre machine du réseau. Ceci permet de décharger la machine dédiée au programme-maître, ce qui améliore généralement la vitesse de réponse de l'interface, de même que les temps de traduction, si le processus-esclave est exécuté sur une machine plus puissante.

Le troisième avantage est, bien sûr, la possibilité d'interrompre une traduction en cours de route (ou toute autre requête dont l'exécution est prise en charge par le processus esclave). Dans une telle implantation, l'écran de base d'IRMA comporte un bouton supplémentaire d'interruption. Lorsque l'usager selectionne ce bouton avec la souris, le processus maître envoie un signal au processus esclave qui a pour effet d'interrompre l'exécution du but courant. Dès lors, le système est prêt à effectuer une autre traduction.

Cette façon de décomposer le travail entre un processus-maître et un processus-esclave (gestion de l'interface versus traitement linguistique) a aussi été adoptée dans l'implantation de l'interface d'utilisation standard de CRITTER [3], et donne d'excellents résultats. Une décomposition plus fine des processus, et en général une utilisation plus globale des possibilités offertes par le parallélisme sont des sujets de recherche qui nous semblent des plus prometteurs.

Conclusion

6

À la suite de son utilisation intensive dans le cadre d'Expotec, période durant laquelle plus de 500 démonstrations ont été effectuées, le kiosque IRMA a été démonté et rapatrié au CCRIT. Dans la même période, il fut décidé de ne pas poursuivre l'entretien et les projets d'amélioration du logiciel qui lui est associé. Les aspects qui ont motivé cette décision sont les suivants:

Premièrement, la quantité de matériel nécessaire au fonctionnement d'IRMA et la complexité de la configuration nous aurait obligé à dédier un emplacement et l'ensemble du matériel aux seules fins de démonstrations d'IRMA. Il s'agit évidemment d'un gaspillage de ressources d'autant plus injustifiable que nous possédons d'autres moyens moins encombrants de démontrer les capacités du système CRITTER.

Deuxièmement, nous considérons avoir atteint un plafond quant au degré d'intégration des périphériques vocaux au système CRITTER. Aller plus loin dans cette entreprise nous obligerait à investir massivement dans l'achat de matériels plus sophistiqués et à envisager des modifications majeures aux techniques développées pour le système de traduction. Or les ressources limitées du groupe de TAO et les objectifs visés pour CRITTER nous forcent pour l'instant à concentrer nos efforts uniquement sur ce dernier projet.

Troisièmement, IRMA a été conçu spécialement pour Expotec 89: la démonstration s'adresse au grand public et met principalement l'accent sur les thèmes de l'exposition. Hors de ce contexte d'utilisation, le système perd une grande partie de sa raison d'être. De plus, en tant qu'outil de démonstration, il ne représente pas adéquatement l'orientation des recherches du groupe de TAO.

La décision d'abandonner le projet IRMA ne sous-entend cependant pas que le bilan de cette expérience soit négatif. Bien au contraire, l'expérience fut enrichissante à plus d'un titre.

IRMA est d'abord et avant tout un logiciel de démonstration. À ce titre, son développement a donné lieu à une réflexion assez poussée sur l'impact de telles démonstrations et sur les principes généraux de leur mise sur pied (cf. chapitre 3). Si un tel exercice n'a, en pratique, pas ou peu de retombées scientifiques directes, son impact comme outil promotionnel et de vulgarisation des résultats d'un projet de recherche fondamentale n'est plus à démontrer. En ce sens, les répercussions d'IRMA sur le projet CRITTER ont été largement au-dessus de nos attentes. Nous avons ainsi acquis la conviction que le développement de logiciels de démonstration devrait figurer dans la planification de tout projet important de recherche informatique au même titre que la rédaction de rapports techniques et d'articles scientifiques.

D'un autre point de vue, IRMA a mis en évidence le potentiel du système CRITTER à participer activement au processus de la reconnaissance vocale. Bien que cette participation dans IRMA fut limitée à une désambiguisation des homophones hétérographes, il apparaît clairement qu'il serait possible d'aller beaucoup plus loin. Ainsi, la mise sur pied d'IRMA nous a donné l'occasion de mesurer l'ampleur et de déterminer la nature des travaux qu'impliquerait l'intégration au système CRITTER d'un périphérique de dictée vocale, advenant que l'implantation d'un tel dispositif s'avère souhaitable dans un contexte d'utilisation réelle d'un progiciel de traduction automatique. Plus précisément, il semble que la solution idéale réside dans l'introduction d'informations phonétiques à même les dictionnaires de CRITTER, et dans l'utilisation d'un système de reconnaissance vocale capable de produire des "treillis de phonèmes". Dans ces conditions, l'utilisation de techniques d'analyse ascendante ("input driven") pour le système de traduction favoriserait le contrôle de la reconnaissance et l'écho visuel pour le locuteur.

Le développement d'un "Dictionnaire Micro-informatisé du Français Phonétique" (DMFP) est présentement envisagé, qui pourrait déboucher sur une augmentation des dictionnaires de CRITTER. Des systèmes de reconnaissance du type mentionné ci-dessus existent, comme le produit "Phonetic Engine", distribué par Speech Systems, Inc. Quant aux méthodes d'analyse, notre groupe explore actuellement la possibilité de remplacer les grammaires actuelles de CRITTER par des mécanismes opérant à partir de la chaîne d'entrée.

Il n'est donc pas impensable ni même irréaliste d'envisager qu'un périphérique de reconnaissance vocale très performant fasse partie intégrante d'un environnemnt d'utilisation d'un système de traduction automatique et ceci dans un avenir assez rapproché.

Bibliographie

[1] P. Isabelle, M. Dymetman, E. Macklovitch, *CRITTER: un système de traduction pour les rapports sur les marchés agricoles*, rapport technique du CCRIT, Ministère des Communications Canada, ISBN 0-662-95181-6, 1988, 24 pages.

[2] L. Bourbeau, F. Pinard, *Dictionnaire Micro-informatisé du Français : Description de ses propriétés morpho-syntaxiques*, Progiciels Bourbeau Pinard inc., ISBN 2-9800622-1-9, 1987, 143 pages.

[3] Jean-Luc Cochard, Pierre Plamondon, *L'interface d'utilisation du système CRITTER*, rapport technique du CCRIT, Ministère des Communications Canada, à paraître, 25 pages.

[4] J.-P. Tubach, *Reconnaissance de parole continue en entrée d'un système de traduction assisté en français et en anglais*, rapport technique du CCRIT, Ministère des Communications Canada, à paraître.

[5] P. Mabilleau, P. Hamel, *Multi_C, A Multitasking Kernel and its Application to Voice Response on a PC*, proceedings of Speech Tech'88, 1988, pp. 157-160.

[6] *Quintus Prolog Reference Manual*, Quintus Computer Systems, Inc., Mountain View, California, 1987.

[7] *Quintus ProWINDOWS User's Guide*, SunView Version, Quintus Computer Systems, Inc. and International Computers, Limited, Mountain View, California, 1988.

[8] *SunView 1 Programmer's Guide*, Sun Microsystems, Inc., 1988.

[9] *SunOS Reference Manual*, Sun Microsystems, Inc., 1988.

[10] *VoiceScribe™-1000: Speech Application Designer's Handbook*, release 2.10, Dragon Systems, MA, 1986.

Annexe

Scénario de démonstration
EXPOTEC 89

Groupe de traduction assistée

Centre canadien de recherches sur
l'informatisation du travail
Ministère des Communications

Laval

juin 1989

Introduction

Ce document décrit en détail un scénario proposé aux animateurs d'EXPOTEC 89 pour les démonstrations du système IRMA. Le scénario en question est conçu de façon à mettre en évidence les caractéristiques les plus intéressantes du système.

Il est important de comprendre que les capacités d'IRMA ne sont nullement limitées au traitement des exemples particuliers utilisés dans ce scénario. Le système est en effet capable de traiter avec une certaine généralité des textes de rapports de marchés agricoles. Le système ne contient pas l'information nécessaire pour traduire quoi que ce soit d'autre. De plus, étant donné qu'IRMA est un prototype de laboratoire qui demeure encore très incomplet, il n'y a aucune garantie de performance adéquate sur des exemples qui seraient "improvisés" par les animateurs. En conséquence il est fortement recommandé que les animateurs s'en tiennent aux exemples qui apparaissent dans le scénario ci-dessous.

La démonstration s'effectue en quatre temps :

- Les trois étapes d'un processus de traduction illustrées par la traduction d'une phrase anglaise,
- la réversibilité d'IRMA illustrée par la retraduction vers l'anglais de la phrase traduite a l'étape précédente,
- la synthèse vocale : traduction d'une phrase anglaise avec haut-parleur branché,
- la reconnaissance vocale : traduction d'une phrase française entrée au microphone.

Avant de commencer :

S'assurer que le système a été démarré correctement et que l'écran affiche le *niveau de départ* de l'interface. Si les zones "texte source" et/ou "texte cible" ne sont pas *vierges*, cliquer sur l'icone poubelle pour les nettoyer.

Premier acte

LES ÉTAPES D'UNE TRADUCTION

Acte I, Scène I : Préliminaires

Commentaire

IRMA est un système *prototype* capable de traduire entre l'anglais et le français, qu'il s'agisse de langue écrite ou orale. Compte tenu du niveau très élevé de difficulté des problèmes, IRMA est orienté vers la traduction d'une sous-langue simple : celle des rapports de marchés agricoles. La technologie est toutefois applicable a beaucoup d'autres sous-domaines.

➡ **Faire apparaître les paramètres de configuration en cliquant (bouton gauche) sur l'icone correspondant du tableau de commandes (au dessus de celui d'IRMA, la vache).**

Commentaire

Dans les étapes ultérieures de la démonstration, nous allons utiliser la reconnaissance vocale. Le processus de reconnaissance utilise des modèles spécifiques à chaque locuteur. Il nous faut donc indiquer au système quel locuteur effectuera l'entrée vocale.

➡ **Vérifier que le bon locuteur est sélectionné et refermer le tableau des paramètres de configuration en déplaçant le pointeur hors de ce tableau.**

➡ **Sélectionner les paramètres suivants:**

langue source : anglais
affichage des structures : oui
microphone : débranché
haut-parleur : débranché

Commentaire

Nous allons maintenant traduire une phrase de l'anglais vers le français. Le processus s'effectuera en trois étapes: l'analyse, le transfert et la synthèse. Au terme de chaque étape, le système affichera des graphiques qui représentent le résultat de son travail.

Acte I, Scène II : Analyse de la phrase

➡ **Faire apparaître le mini-éditeur en cliquant sur l'icone clavier et presser la touche F1. La phrase *prices increased three dollars Monday, remaining steady for the rest of the week* apparaît comme texte source.**

Commentaire

Le système est en mesure de traduire des phrases comme celle-ci, qui est tout-à-fait typique de la sous-langue des rapports de marchés agricoles produits par Agriculture Canada.

➡ **Cliquer sur l'icone IRMA (la vache).**

Commentaire

Observez le changement dans l'icone IRMA pendant que le système est au travail. Observez également que l'icone A (analyse) est assombri. IRMA est en train d'effectuer l'analyse de la phrase. Cette étape correspond à la phase de lecture et compréhension du texte à traduire. Il s'agit d'un processus très complexe, d'une importance cruciale : si le texte en langue source est mal compris, la traduction ne saurait réussir. Une fois l'analyse terminée, le système va construire des graphiques qui décrivent ce qu'il a compris à propos de la phrase analysée.

➡ **Dès que l'icone IRMA revient à la position repos, cliquer sur l'icone A: le graphe sémantique apparaît.**

Commentaire

Ce graphe sémantique constitue une représentation formelle du sens de la phrase analysée. Il rend explicite les objets et les relations que décrit la phrase dans l'univers des rapports de marchés. Par exemple, la notion sémantique de "mouvement ascendant" est formalisée comme une relation "increase(X,Y)" entre une certaine quantité X (dans notre exemple, les prix) et un certain incrément Y (dans notre exemple, un nombre de dollars). En principe, on représente par un même graphe sémantique toutes les phrases qui ont le même sens. Une caractéristique importante des graphes sémantiques est de *rendre explicite* de l'information qui est présente de façon *implicite* dans les phrases. Dans notre phrase exemple, il est relativement explicite que ce qui augmente, ce sont les prix (puisque "prices" figure comme sujet du verbe "increase"). Il est cependant tout-à-fait implicite que ce qui est soutenu ("steady"), ce sont également les prix. Le processus d'analyse découvre ce fait, et l'enregistre dans le graphe sémantique.

➡ **Cliquer sur l'icone GRAM.**

Commentaire

Pour construire le graphe sémantique d'une phrase donnée, IRMA doit d'abord effectuer une décomposition grammaticale complète de cette phrase. Le résultat de cette décomposition est un *arbre syntaxique*. L'arbre de notre phrase exemple indique que nous avons affaire à :

- une phrase principale (*s_princ*) décomposée en une phrase noyau (*s*) et une ponctuation (*pon*);
- la phrase noyau est décomposée en un groupe nominal (*n2*) sujet et un groupe verbal (*v2*);
- le groupe nominal est décomposé en un article (*art*) et un nom commun (*nc*);

- le groupe verbal a pour verbe principal (*vb*) *increased*
- etc.

➡ **Cliquer n'importe où dans la fenêtre "arbre syntaxique" pour retourner à la fenêtre "graphe sémantique".**

Commentaire

Il faut remarquer que le graphe sémantique, bien qu'il représente le sens de la phrase, demeure lié à la langue source. Les concepts qui apparaissent sur les noeuds (comme *increase*) sont orientés vers la langue analysée.

➡ **Cliquer n'importe où dans la fenêtre "graphe sémantique" pour retourner à l'écran principal.**

Acte I, Scène III : Transfert de la phrase

➡ **Cliquer à nouveau sur l'icone IRMA.**

Commentaire

L'étape suivante consiste à produire à partir du graphe sémantique de la phrase analysée un graphe sémantique équivalent, mais orienté vers la langue cible.

➡ **Dès que les icones IRMA et T reviennent à la position repos, cliquer sur l'icone T pour faire apparaître les graphes sémantiques source et cible.**

Commentaire

Dans le graphe cible, les noeuds sont maintenant étiquetés par des concepts orientés vers le vocabulaire de la langue cible. C'est donc durant cette étape de transfert que les opérations de traduction proprement dites sont appliquées. Ces opérations sont grandement simplifiées par le fait qu'elles sont formulées au niveau très *abstrait* des représentations sémantiques. En effet, à ce niveau on fait abstraction des détails grammaticaux qui varient beaucoup d'une langue à l'autre. Dans notre exemple, les graphes source et cible sont parfaitement *isomorphes*, même si les structures grammaticales des phrases concernées sont différentes. Il arrive néanmoins que les graphes sémantiques source et cible soient plus éloignés l'un de l'autre, lorsque les deux langues emploient des moyens très différents pour exprimer la même idée (p.ex. *John swam across the river <--> John a traversé la rivière à la nage*).

➡ **Cliquer n'importe où dans la fenêtre TRANSFERT pour retourner à l'écran principal.**

Acte I, Scène IV: Synthèse de la phrase en langue cible

➡ **Cliquer à nouveau sur l'icone IRMA**

Commentaire

L'icone S (*synthèse*) s'assombrit. IRMA procède maintenant à la synthèse d'une phrase française à partir du graphe sémantique cible. Pour ce faire il faudra consulter un dictionnaire du français et appliquer toutes les règles de la grammaire

française, de façon à ordonner correctement les mots, conjuguer les verbes, effectuer les accords de genre et nombre, etc.

➡ **Dès que les icones IRMA et S reviennent à la position repos, la traduction *les prix augmentaient de 3 dollars lundi, demeurant soutenus pour le reste de la semaine* est affichée sous "texte cible".**

➡ **Cliquer sur l'icone S pour faire apparaître le graphe sémantique cible et la phrase synthétisée.**

Commentaire

IRMA a donc réussi à produire une phrase correcte du français à partir du graphe sémantique cible. Observez que certains des choix qui ont été effectués dans la phrase cible reflètent bien les capacités de compréhension d'IRMA. Par exemple, l'adjectif *soutenus* est correctement accordé avec *les prix* puisque IRMA comprend que ce sont les prix qui sont soutenus.

➡ **Cliquer sur l'icone GRAM.**

Commentaire

Alors que le processus d'analyse part d'une phrase et produit un graphe sémantique en passant par l'intermédiaire d'un arbre syntaxique, le processus de synthèse fait la démarche inverse : produire une phrase à partir d'un graphe sémantique. Encore ici, on passe par l'intermédiaire d'un arbre syntaxique.

➡ **Cliquer n'importe où dans la fenêtre "arbre syntaxique" pour retourner à la fenêtre "graphe sémantique".**

➡ **Cliquer n'importe où dans la fenêtre "graphe sémantique" pour retourner à l'écran principal.**

Deuxième acte
LA RÉVERSIBILITÉ D'IRMA

Commentaire

IRMA possède une particularité unique parmi les systèmes de traduction automatique : il est *réversible*, c'est-à-dire qu'il peut aussi bien traduire du français vers l'anglais que de l'anglais vers le français, en utilisant les mêmes descriptions linguistiques (dictionnaires, grammaires, etc). Afin d'illustrer cette propriété, nous allons retraduire vers l'anglais la phrase qu'IRMA vient de produire. Cette fois nous n'allons pas demander au système d'afficher les graphiques.

➡ **Sélectionner les paramètres suivants :**
langue source : français
affichage des structures : non
microphone : débranché
haut-parleur : débranché

➡ **Facultativement, nettoyer la zone langue source en cliquant sur l'icone poubelle.**

➡ **Faire apparaître le mini-éditeur en cliquant sur l'icone clavier, et presser la touche F1 (la phrase *les prix augmentaient de 3 dollars lundi, demeurant soutenus pour le reste de la semaine.* qui était sous "texte cible" apparaît maintenant comme "texte source").**

➡ **Cliquer sur l'icone IRMA.**

Commentaire

IRMA produit maintenant une retraduction vers l'anglais. (La traduction s'affiche sous "texte cible".) Observez que la traduction obtenue est correcte. Observez également qu'elle n'est pas identique à la phrase anglaise originale qui avait été soumise, mais qu'elle est tout-à-fait équivalente.

➡ **Optionellement, faire réapparaître dans la position texte source cette phrase anglaise originale. À cette fin, mettre le paramètre "langue source" à "anglais", faire apparaître le mini-éditeur et presser F1.**

Commentaire

Cette variation est une conséquence naturelle du fait qu'IRMA connaît généralement de nombreuses façons d'exprimer une même idée. On peut exploiter cette capacité de *paraphrasage* pour mettre un peu de variété dans les traductions obtenues.

Troisième acte

LA SYNTHÈSE VOCALE

Acte III, Scène I

Commentaire

IRMA comprend une composante de synthèse vocale qui permet de prononcer à haute voix les traductions françaises et anglaises qu'il produit.

➡ **Sélectionner les paramètres suivants :**
langue source : anglais
affichage des structures : non
microphone : débranché
haut-parleur : branché

➡ **Faire apparaître le mini-éditeur en cliquant sur l'icone clavier, et presser la touche F2 (la phrase *fat cows were discounted, while the top heifers were up $3-$4* est affichée comme "texte source").**

➡ **Cliquer sur l'icone IRMA.**

Commentaire

Attirer l'attention du spectateur à la fois sur la zone "Texte cible" et sur le haut-parleur. Demander le silence au moment de la synthèse vocale.

➡ **La phrase *les vaches grasses se vendaient à rabais, alors que les meilleures génisses augmentaient de 3 à 4 dollars* s'affiche comme "texte cible" et est prononcée à haute voix par le module de synthèse vocale.**

➡ **Pour réentendre cette synthèse vocale, il suffit de cliquer sur l'icone haut-parleur.**

➡ **Optionnellement, on pourra faire observer qu'IRMA dépasse le stade du mot-à-mot, comme le montre la traduction de *discount* par *vendre à rabais*.**

Acte III, Scène II (optionnelle)

➡ **On peut à ce point montrer que la synthèse se fait également avec l'anglais comme langue cible; pour ce faire, sélectionner le français comme langue source, faire apparaître la phrase source française (F2 dans le mini-éditeur) et effectuer la traduction.**

Quatrième acte

LA RECONNAISSANCE VOCALE

Acte IV, Scène I

Commentaire

IRMA peut également fonctionner par entrée vocale. Compte tenu de l'état de l'art actuel, cette capacité demeure plutôt limitée. La reconnaissance doit s'effectuer par mots détachés et sur la base d'un entraînement spécifique à chaque locuteur.

➡ **Sélectionner les paramètres suivants :**
langue source : français
affichage des structures : non
microphone : branché
haut-parleur : branché

➡ **Faire l'entrée vocale de la phrase *les cours des bovins de court engraissement baissaient de cinq pour cent.***

➡ **Débrancher le microphone.**

Commentaire

Remarquez que le signal vocal contient moins d'information qu'un signal écrit. Le nom *cours* et l'adjectif *court* (ou *courts*) se prononcent de façon identique (ils sont *homophones*). Dans notre phrase exemple, nous avons deux occurrences du son "*cour*", l'une comme nom l'autre comme adjectif. Le nom *bovin* se prononce de la même façon au singulier et au pluriel. Enfin *baissait* et *baissaient* se prononce de la même façon. De plus, la nuance sonore avec *baissé* et *baisser* est trop fine pour que le module de reconnaissance vocale puisse les distinguer de façon fiable.

À la sortie du traitement vocal, nous avons donc beaucoup d'incertitude sur le contenu exact de la phrase à traduire. Pour traduire correctement, IRMA devra être capable de mettre en oeuvre des mécanismes permettant de résoudre ces problèmes d'incertitude.

➡ **Cliquer sur l'icone IRMA, et faire observer la sélection parmi les homophones.**

Commentaire

IRMA ne traduit pas les mots isolément. La traduction, comme nous l'avons vu précédemment, est fondée sur une analyse complète de la grammaire et de la sémantique des phrases. Grâce à la finesse de cette analyse, IRMA a correctement effectué les choix d'homophones dans notre phrase exemple, et nous avons pu obtenir une traduction correcte (*bovin de court engraissement* se traduit par *feeder*).

Cette capacité résulte du niveau d'expertise très élevé d'IRMA en ce qui concerne la grammaire et la sémantique des textes concernés.

Acte IV, Scène II (optionnelle)

➡ **On peut à ce point montrer que la reconnaissance fonctionne également avec l'anglais comme langue source; pour ce faire, sélectionner l'anglais comme langue source, brancher le microphone et faire l'entrée vocale de *lean yearlings sold at steady prices*.**

➡ **Cliquer sur l'icone IRMA, pour effectuer la traduction et la synthèse du français.**